

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 1999, 22:10:49 : Search time 140.58 seconds
(without alignments)
3.355 Million cell updates/sec

Title: US-09-017-524-32
Perfect score: 52
Sequence: TITPCASDAK 10

Scoring table: BLOSUM62

Searched: 142080 seqs, 47160319 residues

Database: PIR-62:1*
1: pir1:1*
2: pir2:1*
3: pir3:1*
4: pir4:1*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	52	100.0	856	1 VCLJH3	env polyprotein pr
2	52	100.0	856	1 VCLJVL	env polyprotein pr
3	52	100.0	841	1 VCLJLV	env polyprotein pr
4	52	100.0	843	1 H44001	env polyprotein pr
5	52	100.0	852	1 VCLJBR	env polyprotein pr
6	52	100.0	856	1 VCLJ3W	env polyprotein pr
7	52	100.0	858	1 VCLJ3W	env polyprotein pr
8	52	100.0	861	1 VCLJ3W	env polyprotein pr
9	52	100.0	861	1 VCLJ3W	env polyprotein pr
10	52	100.0	846	1 VCLJ3W	env polyprotein pr
11	52	100.0	855	1 VCLJ3W	env polyprotein pr
12	52	100.0	859	1 VCLJ3W	env polyprotein pr
13	52	100.0	861	1 VCLJ3W	env polyprotein pr
14	52	100.0	856	1 A44963	env polyprotein pr
15	52	100.0	859	2 T01672	env polyprotein pr
16	52	100.0	506	2 A40218	env polyprotein pr
17	52	100.0	219	2 S25939	env polyprotein pr
18	52	100.0	211	2 S25938	env polyprotein pr
19	52	100.0	495	2 S11493	env polyprotein pr
20	52	100.0	853	2 S54384	env polyprotein pr
21	52	100.0	851	2 S33985	env polyprotein pr
22	52	100.0	854	2 S13288	env polyprotein pr
23	52	100.0	847	2 S13289	env polyprotein pr
24	52	100.0	847	2 T09448	env polyprotein pr
25	52	100.0	853	2 T10314	env polyprotein pr
26	52	100.0	855	2 VCLJ3W	env polyprotein pr
27	52	100.0	854	2 VCLJ3W	env polyprotein pr
28	52	100.0	854	2 VCLJ3W	env polyprotein pr
29	52	100.0	854	2 VCLJ3W	env polyprotein pr
30	52	100.0	854	2 VCLJ3W	env polyprotein pr
31	52	100.0	854	2 VCLJ3W	env polyprotein pr
32	52	100.0	854	2 VCLJ3W	env polyprotein pr
33	52	100.0	854	2 VCLJ3W	env polyprotein pr
34	52	100.0	854	2 VCLJ3W	env polyprotein pr
35	52	100.0	854	2 VCLJ3W	env polyprotein pr
36	52	100.0	854	2 VCLJ3W	env polyprotein pr
37	52	100.0	854	2 VCLJ3W	env polyprotein pr
38	52	100.0	854	2 VCLJ3W	env polyprotein pr
39	52	100.0	854	2 VCLJ3W	env polyprotein pr

40 34 65.4 371 2 S55578 D-nopalline dehydro
41 34 65.4 287 2 C72099 c143 hypobiosal
42 33 63.5 302 2 S52584 c143 protein - BW
43 33 63.5 302 2 A37802 c143 protein - BW
44 33 63.5 677 2 S33608 somatotropin-bind
45 33 63.5 638 2 A33505 somatotropin recep

ALIGNMENTS

RESULT 1
VCLJH3
env polyprotein precursor - human immunodeficiency virus type 1 (isolate HIV-III, BH
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 17-May-1985 #sequence_revision 17-May-1985 #ext-change 16-Jul-1999
C:Accession: A03973
R:Ratner, L.; Haseltine, W.; Paterica, R.; Livak, K.J.; Staudhof, B.; Josephs, S.F.; D
Nberger, J.A.; Papas, T.S.; Chayeh, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.
Nature 313, 277-284, 1985
A:Title: Complete nucleotide sequence of the AIDS virus, HIV-III.
A:Reference number: A93533; M01D:85111123
A:Accession: A03973
A:Molecule type: DNA
A:Residues: 1-856 <RAT>
A:Cross-reference: GB:M15654; GB:K02006; GB:K02007; GB:K02008; GB:K02009; GB:K02010; GB:K02011; GB:K02012; GB:K02013; GB:K02014; GB:K02015; GB:K02016; GB:K02017; GB:K02018; GB:K02019; GB:K02020; GB:K02021; GB:K02022; GB:K02023; GB:K02024; GB:K02025; GB:K02026; GB:K02027; GB:K02028; GB:K02029; GB:K02030; GB:K02031; GB:K02032; GB:K02033; GB:K02034; GB:K02035; GB:K02036; GB:K02037; GB:K02038; GB:K02039; GB:K02040; GB:K02041; GB:K02042; GB:K02043; GB:K02044; GB:K02045; GB:K02046; GB:K02047; GB:K02048; GB:K02049; GB:K02050; GB:K02051; GB:K02052; GB:K02053; GB:K02054; GB:K02055; GB:K02056; GB:K02057; GB:K02058; GB:K02059; GB:K02060; GB:K02061; GB:K02062; GB:K02063; GB:K02064; GB:K02065; GB:K02066; GB:K02067; GB:K02068; GB:K02069; GB:K02070; GB:K02071; GB:K02072; GB:K02073; GB:K02074; GB:K02075; GB:K02076; GB:K02077; GB:K02078; GB:K02079; GB:K02080; GB:K02081; GB:K02082; GB:K02083; GB:K02084; GB:K02085; GB:K02086; GB:K02087; GB:K02088; GB:K02089; GB:K02090; GB:K02091; GB:K02092; GB:K02093; GB:K02094; GB:K02095; GB:K02096; GB:K02097; GB:K02098; GB:K02099; GB:K02100; GB:K02101; GB:K02102; GB:K02103; GB:K02104; GB:K02105; GB:K02106; GB:K02107; GB:K02108; GB:K02109; GB:K02110; GB:K02111; GB:K02112; GB:K02113; GB:K02114; GB:K02115; GB:K02116; GB:K02117; GB:K02118; GB:K02119; GB:K02120; GB:K02121; GB:K02122; GB:K02123; GB:K02124; GB:K02125; GB:K02126; GB:K02127; GB:K02128; GB:K02129; GB:K02130; GB:K02131; GB:K02132; GB:K02133; GB:K02134; GB:K02135; GB:K02136; GB:K02137; GB:K02138; GB:K02139; GB:K02140; GB:K02141; GB:K02142; GB:K02143; GB:K02144; GB:K02145; GB:K02146; GB:K02147; GB:K02148; GB:K02149; GB:K02150; GB:K02151; GB:K02152; GB:K02153; GB:K02154; GB:K02155; GB:K02156; GB:K02157; GB:K02158; GB:K02159; GB:K02160; GB:K02161; GB:K02162; GB:K02163; GB:K02164; GB:K02165; GB:K02166; GB:K02167; GB:K02168; GB:K02169; GB:K02170; GB:K02171; GB:K02172; GB:K02173; GB:K02174; GB:K02175; GB:K02176; GB:K02177; GB:K02178; GB:K02179; GB:K02180; GB:K02181; GB:K02182; GB:K02183; GB:K02184; GB:K02185; GB:K02186; GB:K02187; GB:K02188; GB:K02189; GB:K02190; GB:K02191; GB:K02192; GB:K02193; GB:K02194; GB:K02195; GB:K02196; GB:K02197; GB:K02198; GB:K02199; GB:K02200; GB:K02201; GB:K02202; GB:K02203; GB:K02204; GB:K02205; GB:K02206; GB:K02207; GB:K02208; GB:K02209; GB:K02210; GB:K02211; GB:K02212; GB:K02213; GB:K02214; GB:K02215; GB:K02216; GB:K02217; GB:K02218; GB:K02219; GB:K02220; GB:K02221; GB:K02222; GB:K02223; GB:K02224; GB:K02225; GB:K02226; GB:K02227; GB:K02228; GB:K02229; GB:K02230; GB:K02231; GB:K02232; GB:K02233; GB:K02234; GB:K02235; GB:K02236; GB:K02237; GB:K02238; GB:K02239; GB:K02240; GB:K02241; GB:K02242; GB:K02243; GB:K02244; GB:K02245; GB:K02246; GB:K02247; GB:K02248; GB:K02249; GB:K02250; GB:K02251; GB:K02252; GB:K02253; GB:K02254; GB:K02255; GB:K02256; GB:K02257; GB:K02258; GB:K02259; GB:K02260; GB:K02261; GB:K02262; GB:K02263; GB:K02264; GB:K02265; GB:K02266; GB:K02267; GB:K02268; GB:K02269; GB:K02270; GB:K02271; GB:K02272; GB:K02273; GB:K02274; GB:K02275; GB:K02276; GB:K02277; GB:K02278; GB:K02279; GB:K02280; GB:K02281; GB:K02282; GB:K02283; GB:K02284; GB:K02285; GB:K02286; GB:K02287; GB:K02288; GB:K02289; GB:K02290; GB:K02291; GB:K02292; GB:K02293; GB:K02294; GB:K02295; GB:K02296; GB:K02297; GB:K02298; GB:K02299; GB:K02300; GB:K02301; GB:K02302; GB:K02303; GB:K02304; GB:K02305; GB:K02306; GB:K02307; GB:K02308; GB:K02309; GB:K02310; GB:K02311; GB:K02312; GB:K02313; GB:K02314; GB:K02315; GB:K02316; GB:K02317; GB:K02318; GB:K02319; GB:K02320; GB:K02321; GB:K02322; GB:K02323; GB:K02324; GB:K02325; GB:K02326; GB:K02327; GB:K02328; GB:K02329; GB:K02330; GB:K02331; GB:K02332; GB:K02333; GB:K02334; GB:K02335; GB:K02336; GB:K02337; GB:K02338; GB:K02339; GB:K02340; GB:K02341; GB:K02342; GB:K02343; GB:K02344; GB:K02345; GB:K02346; GB:K02347; GB:K02348; GB:K02349; GB:K02350; GB:K02351; GB:K02352; GB:K02353; GB:K02354; GB:K02355; GB:K02356; GB:K02357; GB:K02358; GB:K02359; GB:K02360; GB:K02361; GB:K02362; GB:K02363; GB:K02364; GB:K02365; GB:K02366; GB:K02367; GB:K02368; GB:K02369; GB:K02370; GB:K02371; GB:K02372; GB:K02373; GB:K02374; GB:K02375; GB:K02376; GB:K02377; GB:K02378; GB:K02379; GB:K02380; GB:K02381; GB:K02382; GB:K02383; GB:K02384; GB:K02385; GB:K02386; GB:K02387; GB:K02388; GB:K02389; GB:K02390; GB:K02391; GB:K02392; GB:K02393; GB:K02394; GB:K02395; GB:K02396; GB:K02397; GB:K02398; GB:K02399; GB:K02400; GB:K02401; GB:K02402; GB:K02403; GB:K02404; GB:K02405; GB:K02406; GB:K02407; GB:K02408; GB:K02409; GB:K02410; GB:K02411; GB:K02412; GB:K02413; GB:K02414; GB:K02415; GB:K02416; GB:K02417; GB:K02418; GB:K02419; GB:K02420; GB:K02421; GB:K02422; GB:K02423; GB:K02424; GB:K02425; GB:K02426; GB:K02427; GB:K02428; GB:K02429; GB:K02430; GB:K02431; GB:K02432; GB:K02433; GB:K02434; GB:K02435; GB:K02436; GB:K02437; GB:K02438; GB:K02439; GB:K02440; GB:K02441; GB:K02442; GB:K02443; GB:K02444; GB:K02445; GB:K02446; GB:K02447; GB:K02448; GB:K02449; GB:K02450; GB:K02451; GB:K02452; GB:K02453; GB:K02454; GB:K02455; GB:K02456; GB:K02457; GB:K02458; GB:K02459; GB:K02460; GB:K02461; GB:K02462; GB:K02463; GB:K02464; GB:K02465; GB:K02466; GB:K02467; GB:K02468; GB:K02469; GB:K02470; GB:K02471; GB:K02472; GB:K02473; GB:K02474; GB:K02475; GB:K02476; GB:K02477; GB:K02478; GB:K02479; GB:K02480; GB:K02481; GB:K02482; GB:K02483; GB:K02484; GB:K02485; GB:K02486; GB:K02487; GB:K02488; GB:K02489; GB:K02490; GB:K02491; GB:K02492; GB:K02493; GB:K02494; GB:K02495; GB:K02496; GB:K02497; GB:K02498; GB:K02499; GB:K02500; GB:K02501; GB:K02502; GB:K02503; GB:K02504; GB:K02505; GB:K02506; GB:K02507; GB:K02508; GB:K02509; GB:K02510; GB:K02511; GB:K02512; GB:K02513; GB:K02514; GB:K02515; GB:K02516; GB:K02517; GB:K02518; GB:K02519; GB:K02520; GB:K02521; GB:K02522; GB:K02523; GB:K02524; GB:K02525; GB:K02526; GB:K02527; GB:K02528; GB:K02529; GB:K02530; GB:K02531; GB:K02532; GB:K02533; GB:K02534; GB:K02535; GB:K02536; GB:K02537; GB:K02538; GB:K02539; GB:K02540; GB:K02541; GB:K02542; GB:K02543; GB:K02544; GB:K02545; GB:K02546; GB:K02547; GB:K02548; GB:K02549; GB:K02550; GB:K02551; GB:K02552; GB:K02553; GB:K02554; GB:K02555; GB:K02556; GB:K02557; GB:K02558; GB:K02559; GB:K02560; GB:K02561; GB:K02562; GB:K02563; GB:K02564; GB:K02565; GB:K02566; GB:K02567; GB:K02568; GB:K02569; GB:K02570; GB:K02571; GB:K02572; GB:K02573; GB:K02574; GB:K02575; GB:K02576; GB:K02577; GB:K02578; GB:K02579; GB:K02580; GB:K02581; GB:K02582; GB:K02583; GB:K02584; GB:K02585; GB:K02586; GB:K02587; GB:K02588; GB:K02589; GB:K02590; GB:K02591; GB:K02592; GB:K02593; GB:K02594; GB:K02595; GB:K02596; GB:K02597; GB:K02598; GB:K02599; GB:K02600; GB:K02601; GB:K02602; GB:K02603; GB:K02604; GB:K02605; GB:K02606; GB:K02607; GB:K02608; GB:K02609; GB:K02610; GB:K02611; GB:K02612; GB:K02613; GB:K02614; GB:K02615; GB:K02616; GB:K02617; GB:K02618; GB:K02619; GB:K02620; GB:K02621; GB:K02622; GB:K02623; GB:K02624; GB:K02625; GB:K02626; GB:K02627; GB:K02628; GB:K02629; GB:K02630; GB:K02631; GB:K02632; GB:K02633; GB:K02634; GB:K02635; GB:K02636; GB:K02637; GB:K02638; GB:K02639; GB:K02640; GB:K02641; GB:K02642; GB:K02643; GB:K02644; GB:K02645; GB:K02646; GB:K02647; GB:K02648; GB:K02649; GB:K02650; GB:K02651; GB:K02652; GB:K02653; GB:K02654; GB:K02655; GB:K02656; GB:K02657; GB:K02658; GB:K02659; GB:K02660; GB:K02661; GB:K02662; GB:K02663; GB:K02664; GB:K02665; GB:K02666; GB:K02667; GB:K02668; GB:K02669; GB:K02670; GB:K02671; GB:K02672; GB:K02673; GB:K02674; GB:K02675; GB:K02676; GB:K02677; GB:K02678; GB:K02679; GB:K02680; GB:K02681; GB:K02682; GB:K02683; GB:K02684; GB:K02685; GB:K02686; GB:K02687; GB:K02688; GB:K02689; GB:K02690; GB:K02691; GB:K02692; GB:K02693; GB:K02694; GB:K02695; GB:K02696; GB:K02697; GB:K02698; GB:K02699; GB:K02700; GB:K02701; GB:K02702; GB:K02703; GB:K02704; GB:K02705; GB:K02706; GB:K02707; GB:K02708; GB:K02709; GB:K02710; GB:K02711; GB:K02712; GB:K02713; GB:K02714; GB:K02715; GB:K02716; GB:K02717; GB:K02718; GB:K02719; GB:K02720; GB:K02721; GB:K02722; GB:K02723; GB:K02724; GB:K02725; GB:K02726; GB:K02727; GB:K02728; GB:K02729; GB:K02730; GB:K02731; GB:K02732; GB:K02733; GB:K02734; GB:K02735; GB:K02736; GB:K02737; GB:K02738; GB:K02739; GB:K02740; GB:K02741; GB:K02742; GB:K02743; GB:K02744; GB:K02745; GB:K02746; GB:K02747; GB:K02748; GB:K02749; GB:K02750; GB:K02751; GB:K02752; GB:K02753; GB:K02754; GB:K02755; GB:K02756; GB:K02757; GB:K02758; GB:K02759; GB:K02760; GB:K02761; GB:K02762; GB:K02763; GB:K02764; GB:K02765; GB:K02766; GB:K02767; GB:K02768; GB:K02769; GB:K02770; GB:K02771; GB:K02772; GB:K02773; GB:K02774; GB:K02775; GB:K02776; GB:K02777; GB:K02778; GB:K02779; GB:K02780; GB:K02781; GB:K02782; GB:K02783; GB:K02784; GB:K02785; GB:K02786; GB:K02787; GB:K02788; GB:K02789; GB:K02790; GB:K02791; GB:K02792; GB:K02793; GB:K02794; GB:K02795; GB:K02796; GB:K02797; GB:K02798; GB:K02799; GB:K02800; GB:K02801; GB:K02802; GB:K02803; GB:K02804; GB:K02805; GB:K02806; GB:K02807; GB:K02808; GB:K02809; GB:K02810; GB:K02811; GB:K02812; GB:K02813; GB:K02814; GB:K02815; GB:K02816; GB:K02817; GB:K02818; GB:K02819; GB:K02820; GB:K02821; GB:K02822; GB:K02823; GB:K02824; GB:K02825; GB:K02826; GB:K02827; GB:K02828; GB:K02829; GB:K02830; GB:K02831; GB:K02832; GB:K02833; GB:K02834; GB:K02835; GB:K02836; GB:K02837; GB:K02838; GB:K02839; GB:K02840; GB:K02841; GB:K02842; GB:K02843; GB:K02844; GB:K02845; GB:K02846; GB:K02847; GB:K02848; GB:K02849; GB:K02850; GB:K02851; GB:K02852; GB:K02853; GB:K02854; GB:K02855; GB:K02856; GB:K02857; GB:K02858; GB:K02859; GB:K02860; GB:K02861; GB:K02862; GB:K02863; GB:K02864; GB:K02865; GB:K02866; GB:K02867; GB:K02868; GB:K02869; GB:K02870; GB:K02871; GB:K02872; GB:K02873; GB:K02874; GB:K02875; GB:K02876; GB:K02877; GB:K02878; GB:K02879; GB:K02880; GB:K02881; GB:K02882; GB:K02883; GB:K02884; GB:K02885; GB:K02886; GB:K02887; GB:K02888; GB:K02889; GB:K02890; GB:K02891; GB:K02892; GB:K02893; GB:K02894; GB:K02895; GB:K02896; GB:K02897; GB:K02898; GB:K02899; GB:K02900; GB:K02901; GB:K02902; GB:K02903; GB:K02904; GB:K02905; GB:K02906; GB:K02907; GB:K02908; GB:K02909; GB:K02910; GB:K02911; GB:K02912; GB:K02913; GB:K02914; GB:K02915; GB:K02916; GB:K02917; GB:K02918; GB:K02919; GB:K02920; GB:K02921; GB:K02922; GB:K02923; GB:K02924; GB:K02925; GB:K02926; GB:K02927; GB:K02928; GB:K02929; GB:K02930; GB:K02931; GB:K02932; GB:K02933; GB:K02934; GB:K02935; GB:K02936; GB:K02937; GB:K02938; GB:K02939; GB:K02940; GB:K02941; GB:K02942; GB:K02943; GB:K02944; GB:K02945; GB:K02946; GB:K02947; GB:K02948; GB:K02949; GB:K02950; GB:K02951; GB:K02952; GB:K02953; GB:K02954; GB:K02955; GB:K02956; GB:K02957; GB:K02958; GB:K02959; GB:K02960; GB:K02961; GB:K02962; GB:K02963; GB:K02964; GB:K02965; GB:K02966; GB:K02967; GB:K02968; GB:K02969; GB:K02970; GB:K02971; GB:K02972; GB:K02973; GB:K02974; GB:K02975; GB:K02976; GB


```

OY      1  TITFCASDAK 10
Db      49  TITFCASDAK 58

RESULT  7
VCLJH4
env polypeptide - human immunodeficiency virus type 1 (isolate CDC-451)
N:Alternate names: coat polypeptide
N:Contains: coat protein gp120; coat protein gp41
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C:Accession: C25523
R:Resat S M., Kalyanaram, V S., Casey, J M., Sriinivasan, A., Andersen, P R., Devare,
Proc. Natl. Acad. Sci. U S A. 83, 8380-8384, 1986
A:Title: Molecular cloning and primary nucleotide sequence analysis of a distinct human
A:Reference number: A42995, M01D:92351552
A:Accession: C25523
A:Molecule type: DNA
A:Residues: 1-864 <DES>
A:Cross-references: GR:M1917, NID:9326460, P1DN AAA4411.1; PID:9326467
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polypeptide
C:Keywords: capsid protein; coat protein; glycoprotein; polypeptide; transmembrane prot
F:1-521/Product: coat protein gp120 #status predicted <SPI>
F:522-869/Product: coat protein gp41 #status predicted <GPI>
F:870-1339/Product: coat protein gp120 #status predicted <SPI>
F:870-1339,142-165,195-198,208-245,252-273,287-300,312,342,349,365,371,395,405,409,459

Query Match      100.0%  Score 52: DB 1: Length 864:
Best local similarity 100.0%  Pred. No. 0.024:
Matches 10: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY      1  TITFCASDAK 10
Db      51  TITFCASDAK 60

RESULT  8
VCLJKB
env polypeptide precursor - human immunodeficiency virus type 1 (strain KB-1-gp41)
N:Alternate names: coat polypeptide
N:Contains: coat protein gp120; coat protein gp41
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 01-Mar-1996
C:Accession: A42995
R:Shimizu, H., Hasebe, F., Tsuchie, H., Morikawa, S., Ushijima, H., Kitamura, T.
Virology 189, 534-546, 1992
A:Title: Analysis of a human immunodeficiency virus type 1 isolate carrying a truncated
A:Reference number: A42995, M01D:92351552
A:Accession: A42995
A:Molecule type: mRNA
A:Residues: 1-961 <SH1>
A:Cross-references: GR:S41265, GB:E01206
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polypeptide
C:Keywords: coat protein; glycoprotein; polypeptide; transmembrane protein
F:1-689/Domain: extracellular #status predicted <EXT>
F:1-33/Domain: signal sequence #status predicted <SIG>
F:34-517/Product: coat protein gp120 #status predicted <SPI>
F:518-517/Product: cleavage processing #status predicted <CP1>
F:514-517/Region: cleavage processing #status predicted <CP2>
F:518-861/Region: hydrophobic #status predicted <HYD>
F:518-334/Region: hydrophobic #status predicted <CP2>
F:690-711/Domain: transmembrane #status predicted <TM1>
F:712-861/Domain: intracellular #status predicted <INT>
F:756-772/Region: hydrophobic #status predicted <HYD>
F:93,141,145,146,163,191,192,237,241,246,269,283,296,308,338,345,361,367,397,403,408,414

```

```

Query Match      100.0%  Score 52: DB 1: Length 861:
Best local similarity 100.0%  Pred. No. 0.023:
Matches 10: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY      1  TITFCASDAK 10
Db      55  TITFCASDAK 64

RESULT  9
VCLJXK
env polypeptide precursor - human immunodeficiency virus type 1 (strain KB-1-gp22)
N:Alternate names: coat polypeptide
N:Contains: coat protein gp120; coat protein gp41
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 01-Mar-1996
C:Accession: B42995
R:Shimizu, H., Hasebe, F., Tsuchie, H., Morikawa, S., Ushijima, H., Kitamura, T.
Virology 189, 534-546, 1992
A:Title: Analysis of a human immunodeficiency virus type 1 isolate carrying a truncated
A:Reference number: A42995, M01D:92351552
A:Accession: B42995
A:Molecule type: mRNA
A:Residues: 1-729 <SH1>
A:Cross-references: GB:S41265, GB:E01206
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polypeptide
C:Keywords: coat protein; glycoprotein; polypeptide; transmembrane protein
F:1-689/Domain: extracellular #status predicted <EXT>
F:1-33/Domain: signal sequence #status predicted <SIG>
F:34-517/Product: coat protein gp120 #status predicted <SPI>
F:514-517/Region: cleavage processing #status predicted <CP1>
F:518-517/Product: cleavage processing #status predicted <CP2>
F:518-729/Product: coat protein gp41 #status predicted <SPI>
F:518-534/Region: hydrophobic #status predicted <HYD>
F:690-711/Domain: transmembrane #status predicted <TM1>
F:712-729/Domain: intracellular #status predicted <INT>
F:93,141,145,146,163,191,192,237,241,246,269,283,296,308,338,345,361,367,397,403,408,

Query Match      100.0%  Score 52: DB 1: Length 729:
Best local similarity 100.0%  Pred. No. 0.02:
Matches 10: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY      1  TITFCASDAK 10
Db      55  TITFCASDAK 64

RESULT  10
VCLJND
env polypeptide precursor - human immunodeficiency virus type 1 (isolate NTK)
N:Alternate names: coat polypeptide
N:Contains: coat protein gp120; coat protein gp41
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1994
C:Accession: J00066
R:Spiller, B., Sire, C., Zachar, V., Key, F., Baire-Ehrmann, F., Gailbault, P., Haege,
Gene 81, 275-284, 1989
A:Title: Nucleotide sequence of HIV-1 NTK, a highly cytopathic strain of the human im
A:Reference number: J00065, M01D:9094200
A:Accession: J00066
A:Molecule type: DNA
A:Residues: 1-846 <SPI>
A:Cross-references: GR:M2723, NID:9229154, P1DN AAA4411.1; PID:9229154
A:Note: the authors translated the codon 323 for residue 523 as Arg
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polypeptide

```


Query Match 100.0%; Score 52; DB 1; Length 856;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TITFCASDAK 10
|||||
Db 49 TITFCASDAK 58

RESULT 15

T01672
envelope polypeptide precursor - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 05-Mar-1999
C:Accession: T01672
R:Alizon, M.; Main-Hobson, S.; Gluckman, J.C.; Sonigo, P.
Cell 46, 63-74, 1986
A:Title: Genetic variability of the AIDS virus: Nucleotide sequence analysis of two isol
A:Reference number: Z14389; MUID:85245056
A:Accession: T01672
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-859<ALL>
A:Cross-References: EMBL:R03456; NID:g328018; PID:g328026
C:Superfamily: type E retrovirus env polypeptide

Query Match 100.0%; Score 52; DB 2; Length 859;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TITFCASDAK 10
|||||
Db 49 TITFCASDAK 58

Search completed: December 9, 1999, 22:10:49
Job time: 1510 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 1999, 22:12:23 : Search time 84.36 seconds

(without alignments)
3.448 Million cell updates/sec

Title: US-09-017-524-32

Sequence: 1 TTLFCASDAK 10

Scoring table: BLOSUM62

Searched: 80000 seqs, 29085965 residues

Database: SwissProt_38:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	856	1	ENV_HV1B1
2	52	100.0	851	1	ENV_HV1B8
3	52	100.0	852	1	ENV_HV1B1
4	52	100.0	861	1	ENV_HV1B1
5	52	100.0	858	1	ENV_HV1C4
6	52	100.0	853	1	ENV_HV1E1
7	52	100.0	854	1	ENV_HV1H2
8	52	100.0	856	1	ENV_HV1H3
9	52	100.0	867	1	ENV_HV1J3
10	52	100.0	848	1	ENV_HV1J8
11	52	100.0	861	1	ENV_HV1K8
12	52	100.0	856	1	ENV_HV1LW
13	52	100.0	859	1	ENV_HV1LM
14	52	100.0	853	1	ENV_HV1MF
15	52	100.0	856	1	ENV_HV1N5
16	52	100.0	421	1	ENV_HV1N5
17	52	100.0	846	1	ENV_HV1ND
18	52	100.0	856	1	ENV_HV1PV
19	52	100.0	847	1	ENV_HV1S1
20	52	100.0	852	1	ENV_HV1S3
21	52	100.0	856	1	ENV_HV1SC
22	52	100.0	856	1	ENV_HV1W1
23	52	100.0	847	1	ENV_HV1W2
24	52	100.0	843	1	ENV_HV1Y2
25	52	100.0	853	1	ENV_HV1Z2
26	52	100.0	460	1	ENV_HV1Z3
27	52	100.0	855	1	ENV_HV1Z6
28	52	100.0	863	1	ENV_HV1Z8
29	52	100.0	856	1	ENV_HV1Z8
30	49	94.2	855	1	ENV_HV1A2
31	49	94.2	855	1	ENV_HV1OY
32	48	92.3	865	1	ENV_HV1RH
33	42	80.8	854	1	ENV_HV1CZ
34	40	76.9	1476	1	A2M1_MOUSE
35	39	75.0	1451	1	A2M2_MOUSE
36	36	69.2	329	1	YH03_YEAST
37	35	67.3	515	1	PEPB_MYCTU
38	35	67.3	553	1	SVU_ECOLI
39	34	65.4	1477	1	AL13_RAT
40	33	63.5	302	1	PTFE_FPMWP
41	33	63.5	550	1	GHRH_MOUSE
42	33	63.5	638	1	GHR_RAT
43	32	61.5	287	1	Y552_HELPY

ALIGNMENTS

RESULT	ID	ENV_HV1B1	STANDARD:	PRT:	856 AA.
AC	P03375:				
DT	21-JUL-1986 (Rel. 01, Created)				
DT	21-JUL-1986 (Rel. 01, Last sequence update)				
DT	15-JUL-1999 (Rel. 38, Last annotation update)				
DE	ENVELOPE POLYPEPTIDE GP160 PRECURSOR (CONTAINS: EXTERIOR MEMBRANE GLYCOPROTEIN (GP120), TRANSMEMBRANE GLYCOPROTEIN (GP41)).				
GN	ENV.				
OS	Human immunodeficiency virus type 1 (HIV isolate) (HIV-1).				
CC	Viruses; Retroviral viruses; Retroviridae; Lentivirus.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 8511123.				
RA	RATNER L., HASELTINE W., PATARCA P., LIVAK K.J., STAROCH B.P., JOSEPHS S.F., DORAN E.R., RAFALESKI J.A., WHITEHORN E.A., BAUMEISTER K., IVANOFF L., PETTEWAY S.P., JR., PEARSON M.L., LAUTENBERGER J.A., PAPAS T.S., GRABER J., CHANG N.T., GALLO R.C., WONG-STAL F.,				
RT	"Complete nucleotide sequence of the AIDS virus, HIV-III."				
RL	Nature 313:277-284(1985).				
RN	[2]				
RP	DISULFIDE BONDS, ANTICARBOHYDRATE-BINDING SITES.				
RX	MEDLINE; 90285159.				
RA	LEONARD C.K., SPELTMAN M.W., FLOTTER L., HARRIS P.J., THOMAS J.N., GREGORY T.J.,				
RT	"Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells."				
RL	J Biol. Chem. 265:10473-10482(1990).				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/infocenter/ or send an email to license@sib.ch).				
CC	EMBL; M15654; AAA44205.1; -				
DR	PIR; A03973; VCLJH3.				
DR	HIV; M15654; ENVSBH102.				
DR	PFAM; PF00516; GP120.1.				
DR	PFAM; PF00517; GP41.1.				
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.				
FT	SIGNAL	1	30		
FT	CHAIN	31	511		EXTERIOR MEMBRANE GLYCOPROTEIN
FT	CHAIN	512	856		TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	54	74		
FT	DISULFID	119	205		
FT	DISULFID	126	196		
FT	DISULFID	131	157		
FT	DISULFID	218	247		
FT	DISULFID	228	239		
FT	DISULFID	296	331		
FT	DISULFID	378	445		
FT	DISULFID	385	418		
FT	CARBOHYD	88	88		
FT	CARBOHYD	136	136		
FT	CARBOHYD	141	141		
FT	CARBOHYD	156	156		
FT	CARBOHYD	160	160		
FT	CARBOHYD	186	186		

```

FT CARBOHYD 197 197
FT CARBOHYD 230 230
FT CARBOHYD 234 234
FT CARBOHYD 241 241
FT CARBOHYD 262 262
FT CARBOHYD 276 276
FT CARBOHYD 289 289
FT CARBOHYD 295 295
FT CARBOHYD 301 301
FT CARBOHYD 332 332
FT CARBOHYD 339 339
FT CARBOHYD 356 356
FT CARBOHYD 386 386
FT CARBOHYD 392 392
FT CARBOHYD 397 397
FT CARBOHYD 406 406
FT CARBOHYD 448 448
FT CARBOHYD 463 463
FT CARBOHYD 511 511
FT CARBOHYD 519 519
FT CARBOHYD 625 625
FT CARBOHYD 637 637
FT CARBOHYD 674 674
FT CARBOHYD 750 750
FT CARBOHYD 816 816
SQ SEQUENCE 856 AA: 97224 MW: 132PAREB CRC32:

```

```

Query Match 100.0% Score 52; DB 1; Length 856;
Best Local Similarity 100.0%; Pred. No. 0.009;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TITFCASDAK 10
DB 50 TITFCASDAK 59

```

```

RESULT 2
ENV_HV1BR STANDARD: PRT: 851 AA.
ID ENV_HV1BR
AC P04582:
DI 13-AUG-1987 (rel. 05, Last sequence update)
DI 15-JUL-1999 (rel. 38, Last annotation update)
DE ENVELOPE POLYPEPTIDE GP160 PRECURSOR (CONTAINS: EXTERIOR MEMBRANE
  GLYCOPROTEIN (GP120); TRANSMEMBRANE GLYCOPROTEIN (GP41)).
GN ENV.
OS Human immunodeficiency virus type 1 (BR8 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 8511123.
RA RATNER L., HASSETTINE W., PATARCA R., LIVAK K.J., STARCH B.R.,
  JOSEPH S.F., DORAN F.P., PAFALSKI J.A., WHITEORN E.A.,
  RADWISZTER K., IVANOFF L., PETTEWAY S.R., JR., PEARSON M.L.,
  LAUTENBERGER J.A., PAPAS T.S., GRAYNER J., CHANG N.T., GALLO R.C.,
  WONG STALL F.:
  "Complete nucleotide sequence of the AIDS virus, HTLV-III."
  Nature 313:277-284(1985).

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation
  at the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).

```

```

DB EMBL: K02011: AAA44661.1: -
DB HIV: K02011: ENVSHB8.
DB PFAM: PF00515: GP120: 1.
DB PFAM: PF00517: GP41: 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;

```

```

KW Signal. 1 30
FT CHAIN 31 506
FT CHAIN 507 851
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 440
FT DISULFID 385 413
FT CARBOHYD 88 88
FT CARBOHYD 136 136
FT CARBOHYD 141 141
FT CARBOHYD 156 156
FT CARBOHYD 160 160
FT CARBOHYD 186 186
FT CARBOHYD 197 197
FT CARBOHYD 230 230
FT CARBOHYD 234 234
FT CARBOHYD 241 241
FT CARBOHYD 262 262
FT CARBOHYD 276 276
FT CARBOHYD 295 295
FT CARBOHYD 301 301
FT CARBOHYD 332 332
FT CARBOHYD 339 339
FT CARBOHYD 356 356
FT CARBOHYD 386 386
FT CARBOHYD 392 392
FT CARBOHYD 401 401
FT CARBOHYD 443 443
FT CARBOHYD 458 458
FT CARBOHYD 606 606
FT CARBOHYD 611 611
FT CARBOHYD 620 620
FT CARBOHYD 632 632
FT CARBOHYD 669 669
FT CARBOHYD 745 745
FT CARBOHYD 811 811
SQ SEQUENCE 851 AA: 96644 MW: 14224E47 CRC32:

```

```

Query Match 100.0% Score 52; DB 1; Length 851;
Best Local Similarity 100.0%; Pred. No. 0.009;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TITFCASDAK 10
DB 50 TITFCASDAK 59

```

```

RESULT 3
ENV_HV1BN STANDARD: PRT: 852 AA.
ID ENV_HV1BN
AC P12488:
DI 01-OCT-1989 (rel. 12, Created)
DI 01-OCT-1989 (rel. 12, Last sequence update)
DI 15-JUL-1999 (rel. 38, Last annotation update)
DE ENVELOPE POLYPEPTIDE GP160 PRECURSOR (CONTAINS: EXTERIOR MEMBRANE
  GLYCOPROTEIN (GP120); TRANSMEMBRANE GLYCOPROTEIN (GP41)).
GN ENV.
OS Human immunodeficiency virus type 1 (BR8 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 89085613.
RA ANAND R., THAYER R., SPINIVASAN A., NAVYAP S., GARDNER M., LUCI W.P.,
  DANDEKAR S.:
  "Biological and molecular characterization of human immunodeficiency
  virus (HIV-1BR) from the brain of a patient with progressive

```


Query Match 100.0% Score 52; DB 1; Length 861;
 Best Local Similarity 100.0%; Pred. No. 0.0091;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TITFCASDAK 10
 |||||
 Db 50 TITFCASDAK 59

RESULT 5
 ID ENV_HV1C4 STANDARD; PRT; 868 AA.
 AC P05879;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ENVELOPE POLYPEPTIDE GP160 PRECURSOR [CONTAINS: EXTERIOR MEMBRANE
 DE GLYCOPROTEIN (GP120); TRANSMEMBRANE GLYCOPROTEIN (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (CDC-45) isolate (HIV-1).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 87041461.
 RA DESAI S.M., KALYANAPAMAN V.S., CASEY T.M., SPINIVASAN A.,
 RA ANDERSEN P.R., DEVAE S.G.;
 RT "Molecular cloning and primary nucleotide sequence analysis of a
 RT distinct human immunodeficiency virus isolate reveal significant
 RT divergence in its genomic sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:8380-8384(1986).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; M13137; AAA44311.1;
 DR PIR; G25523; VCLH4.
 DR HIV; M13137; ENVSOC45;
 DR PFAM; PF00516; GP120; 1.
 DR AIDS; Coat Protein; Polyprotein; Glycoprotein; Transmembrane;
 KW Signal.
 FT SIGNAL. 1 32
 FT CHAIN 33 522 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 523 868 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 55 75 BY SIMILARITY.
 FT DISULFID 120 216 BY SIMILARITY.
 FT DISULFID 127 207 BY SIMILARITY.
 FT DISULFID 132 163 BY SIMILARITY.
 FT DISULFID 229 258 BY SIMILARITY.
 FT DISULFID 249 250 BY SIMILARITY.
 FT DISULFID 307 341 BY SIMILARITY.
 FT DISULFID 387 456 BY SIMILARITY.
 FT DISULFID 394 429 BY SIMILARITY.
 FT CARBOHYD 89 89 POTENTIAL.
 FT CARBOHYD 131 131 POTENTIAL.
 FT CARBOHYD 138 138 POTENTIAL.
 FT CARBOHYD 139 139 POTENTIAL.
 FT CARBOHYD 142 142 POTENTIAL.
 FT CARBOHYD 162 162 POTENTIAL.

Query Match 100.0% Score 52; DB 1; Length 866;
 Best Local Similarity 100.0%; Pred. No. 0.0091;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TITFCASDAK 10
 |||||
 Db 51 TITFCASDAK 60

RESULT 6
 ID ENV_HV1E1 STANDARD; PRT; 853 AA.
 AC P04581;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ENVELOPE POLYPEPTIDE GP160 PRECURSOR [CONTAINS: EXTERIOR MEMBRANE
 DE GLYCOPROTEIN (GP120); TRANSMEMBRANE GLYCOPROTEIN (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (ELI isolate) (HIV-1).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86245056.
 RA ALIZON M., MAIN-HOBSON S., MONTANIER I., SONTGO P.;
 RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
 RT of two isolates from African patients.";
 RL Cell 46:63-74(1986).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; K03454; AAA44329.1;
 DR EMBL; A07108; CAA00616.1;
 DR HIV; K03454; ENVSSEL.
 DR PFAM; PF00516; GP120; 1.
 DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
 KW Signal.
 FT SIGNAL. 1 29

```

FT CHAIN 30 508 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 509 853 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 206 BY SIMILARITY.
FT DISULFID 125 197 BY SIMILARITY.
FT DISULFID 130 154 BY SIMILARITY.
FT DISULFID 215 248 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 297 230 BY SIMILARITY.
FT DISULFID 376 412 BY SIMILARITY.
FT DISULFID 484 416 BY SIMILARITY.
FT CARBOHYD 87 87 POTENTIAL.
FT CARBOHYD 137 137 POTENTIAL.
FT CARBOHYD 143 143 POTENTIAL.
FT CARBOHYD 153 153 POTENTIAL.
FT CARBOHYD 157 157 POTENTIAL.
FT CARBOHYD 183 183 POTENTIAL.
FT CARBOHYD 188 188 POTENTIAL.
FT CARBOHYD 198 198 POTENTIAL.
FT CARBOHYD 235 235 POTENTIAL.
FT CARBOHYD 242 242 POTENTIAL.
FT CARBOHYD 263 263 POTENTIAL.
FT CARBOHYD 277 277 POTENTIAL.
FT CARBOHYD 290 290 POTENTIAL.
FT CARBOHYD 331 331 POTENTIAL.
FT CARBOHYD 353 353 POTENTIAL.
FT CARBOHYD 384 384 POTENTIAL.
FT CARBOHYD 390 390 POTENTIAL.
FT CARBOHYD 394 394 POTENTIAL.
FT CARBOHYD 400 400 POTENTIAL.
FT CARBOHYD 405 405 POTENTIAL.
FT CARBOHYD 406 406 POTENTIAL.
FT CARBOHYD 411 411 POTENTIAL.
FT CARBOHYD 445 445 POTENTIAL.
FT CARBOHYD 458 458 POTENTIAL.
FT CARBOHYD 459 459 POTENTIAL.
FT CARBOHYD 462 462 POTENTIAL.
FT CARBOHYD 608 608 POTENTIAL.
FT CARBOHYD 613 613 POTENTIAL.
FT CARBOHYD 622 622 POTENTIAL.
FT CARBOHYD 634 634 POTENTIAL.
SO SEQUENCE 853 AA: 96721 MW: 90562P8A CPG12:

Query Match 100.0% Score 52: DB 1: Length 853;
Best Local Similarity 100.0%: Pred No. 0.009;
Matches 10: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TITFCASDAK 10
Db 49 TITFCASDAK 58

```

```

RESULT 7
ENV_HV1H2 STANDARD: PRT: 856 AA.
ID ENV_HV1H2
AC P04578:
DI 13-AUG-1987 (Rel. 05, Created)
DI 15-JUN-1999 (Rel. 38, Last sequence update)
DI 15-JUL-1999 (Rel. 38, Last annotation update)
DE ENVELOPE POLYPEPTIDE gp160 PRECURSOR [CONTAINS: EXTERIOR MEMBRANE
GLYCOPROTEIN (GP120); TRANSMEMBRANE GLYCOPROTEIN (GP41)].
ENV.
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
OC Viruses: Retroviridae; Lentivirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE, 8729336.
RA RATNER L., FISHER A., JAGGRTINSKI L.L., MITSUYA H., LIU P.-S.,
GALEO R.C., WONG-STAL F.:
RT "Complete nucleotide sequences of functional clones of the AIDS
virus."

```

```

RL AIDS Res. Hum. Retroviruses 3:57-69(1987).
RN [3]
RP REVISIONS.
RA RATNER L., FISHER A., JAGGRTINSKI L.L., MITSUYA H., LIU P.-S.,
RA GALEO R.C., WONG-STAL F.:
RA Submitted (APR-1997) to the EMBL/Genbank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.ebi.ac.uk/infocentre
CC or send an email to license@sib.ac.uk.)
CC -----
CC EMBL, K03455; AB50262.1; -.
DR HIV; K03455; ENVSHXB2.
DR PFAM; PF00516; GP120; 1.
DR PFAM; PF00517; GP41; 1.
KW AIDS; Coat protein, Polypeptide, Glycoprotein, Transmembrane;
KW Signal.
FT CHAIN 1 30
FT CHAIN 512 512 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 119 205 BY SIMILARITY.
FT DISULFID 126 195 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 331 BY SIMILARITY.
FT DISULFID 378 445 BY SIMILARITY.
FT DISULFID 385 418 BY SIMILARITY.
FT CARBOHYD 88 88 POTENTIAL.
FT CARBOHYD 136 136 POTENTIAL.
FT CARBOHYD 141 141 POTENTIAL.
FT CARBOHYD 156 156 POTENTIAL.
FT CARBOHYD 160 160 POTENTIAL.
FT CARBOHYD 186 186 POTENTIAL.
FT CARBOHYD 197 197 POTENTIAL.
FT CARBOHYD 230 230 POTENTIAL.
FT CARBOHYD 234 234 POTENTIAL.
FT CARBOHYD 241 241 POTENTIAL.
FT CARBOHYD 262 262 POTENTIAL.
FT CARBOHYD 276 276 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
FT CARBOHYD 295 295 POTENTIAL.
FT CARBOHYD 301 301 POTENTIAL.
FT CARBOHYD 332 332 POTENTIAL.
FT CARBOHYD 339 339 POTENTIAL.
FT CARBOHYD 356 356 POTENTIAL.
FT CARBOHYD 356 356 POTENTIAL.
FT CARBOHYD 386 386 POTENTIAL.
FT CARBOHYD 392 392 POTENTIAL.
FT CARBOHYD 397 397 POTENTIAL.
FT CARBOHYD 406 406 POTENTIAL.
FT CARBOHYD 448 448 POTENTIAL.
FT CARBOHYD 463 463 POTENTIAL.
FT CARBOHYD 611 611 POTENTIAL.
FT CARBOHYD 616 616 POTENTIAL.
FT CARBOHYD 624 624 POTENTIAL.
FT CARBOHYD 637 637 POTENTIAL.
FT CARBOHYD 674 674 POTENTIAL.
FT CARBOHYD 750 750 POTENTIAL.
FT CARBOHYD 816 816 POTENTIAL.
SO SEQUENCE 856 AA: 97212 MW: F0436A03 CPG22:

```

```

Query Match 100.0% Score 52: DB 1: Length 856;
Best Local Similarity 100.0%: Pred No. 0.009;
Matches 10: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TITFCASDAK 10

```

Db 50 TTLPCASDAK 59

RESULT 8

ENV_HVL13 STANDARD: PRT: 856 AA.

AC P04524:

DT 13-AUG-1987 (Rel. 05, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE ENVELOPE POLYPEPTIDE GP160 PRECURSOR [CONTAINS: EXTERIOR MEMBRANE GLYCOPROTEIN (GP120); TRANSMEMBRANE GLYCOPROTEIN (GP41)]

GN ENV.

OS Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).

OC Viruses; Retroviroidea; Retroviridae; Lentivirus.

RA [1]

RA MEDLINE: 85228248.

RA SHAW G.M., WONG-STAL F., REDDY E.P.,

RT "HIV-111 env gene products synthesized in E. coli are recognized by RT antibodies present in the sera of AIDS patients.";

RL (Cell 41:979-986(1985)).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M14100; AAA44679.1; -

DR HIV: M14100; ENVSHX3.

DR PFAM: PF00516; GP120; 1

DR PFAM: PF00517; GP41; 1

KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.

FT SIGNAL 1 40

FT CHAIN 31 511

FT DISULFID 119 205

FT DISULFID 126 196

FT DISULFID 131 157

FT DISULFID 218 247

FT DISULFID 228 239

FT DISULFID 296 331

FT DISULFID 328 445

FT DISULFID 485 418

FT CARBOHYD 88 88

FT CARBOHYD 136 136

FT CARBOHYD 141 141

FT CARBOHYD 156 156

FT CARBOHYD 160 160

FT CARBOHYD 186 186

FT CARBOHYD 197 197

FT CARBOHYD 230 230

FT CARBOHYD 234 234

FT CARBOHYD 241 241

FT CARBOHYD 252 252

FT CARBOHYD 276 276

FT CARBOHYD 289 289

FT CARBOHYD 295 295

FT CARBOHYD 301 301

FT CARBOHYD 332 332

FT CARBOHYD 339 339

FT CARBOHYD 356 356

FT CARBOHYD 386 386

FT CARBOHYD 392 392

FT CARBOHYD 397 397

FT CARBOHYD 406 406

FT CARBOHYD 448 448

FT CARBOHYD 463 463

EXTERIOR MEMBRANE GLYCOPROTEIN

TRANSMEMBRANE GLYCOPROTEIN

Query Match 100.0%; Score 52; DB 1; Length 856;

Best Local Similarity 100.0%; Pctd. No. n nna;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 TTLPCASDAK 10

Db 50 TTLPCASDAK 59

RESULT 9

ENV_HVL13 STANDARD: PRT: 867 AA.

AC P12489:

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE ENVELOPE POLYPEPTIDE GP160 PRECURSOR [CONTAINS: EXTERIOR MEMBRANE GLYCOPROTEIN (GP120); TRANSMEMBRANE GLYCOPROTEIN (GP41)]

GN ENV.

OS Human immunodeficiency virus type 1 (JH3 isolate) (HIV-1).

OC Viruses; Retroviroidea; Retroviridae; Lentivirus.

RA [1]

RA MEDLINE: 89352108.

RA KOMIYAMA N., HATTORI N., IMURA J., SAKUMA S., KURIHARA T., YOSHIDA M.,

RT "Nucleotide sequences of gag and env genes of a Japanese isolate of HIV-1 and their expression in bacteria.";

RL AIDS Res. Hum. Retroviruses 5:411-419(1989).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M21138; AAB03526.1; -

DR HIV: M21138; ENVSHX3.

DR PFAM: PF00516; GP120; 1

DR PFAM: PF00517; GP41; 1

KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.

FT SIGNAL 1 30

FT CHAIN 31 516

FT DISULFID 53 73

FT DISULFID 118 217

FT DISULFID 125 208

FT DISULFID 130 160

FT DISULFID 230 259

FT DISULFID 240 251

FT DISULFID 308 342

FT DISULFID 388 457

FT DISULFID 395 430

FT CARBOHYD 87 87

FT CARBOHYD 135 135

FT CARBOHYD 140 140

FT CARBOHYD 143 143

FT CARBOHYD 159 159

FT CARBOHYD 163 163

FT CARBOHYD 188 188

FT CARBOHYD 189 189

EXTERIOR MEMBRANE GLYCOPROTEIN

TRANSMEMBRANE GLYCOPROTEIN

FT CARBOHYD 199 199 POTENTIAL.
 FT CARBOHYD 209 209 POTENTIAL.
 FT CARBOHYD 246 246 POTENTIAL.
 FT CARBOHYD 253 253 POTENTIAL.
 FT CARBOHYD 274 274 POTENTIAL.
 FT CARBOHYD 288 288 POTENTIAL.
 FT CARBOHYD 307 307 POTENTIAL.
 FT CARBOHYD 350 350 POTENTIAL.
 FT CARBOHYD 355 355 POTENTIAL.
 FT CARBOHYD 372 372 POTENTIAL.
 FT CARBOHYD 396 396 POTENTIAL.
 FT CARBOHYD 402 402 POTENTIAL.
 FT CARBOHYD 408 408 POTENTIAL.
 FT CARBOHYD 412 412 POTENTIAL.
 FT CARBOHYD 418 418 POTENTIAL.
 FT CARBOHYD 423 423 POTENTIAL.
 FT CARBOHYD 450 450 POTENTIAL.
 FT CARBOHYD 475 475 POTENTIAL.
 FT CARBOHYD 622 622 POTENTIAL.
 FT CARBOHYD 627 627 POTENTIAL.
 FT CARBOHYD 636 636 POTENTIAL.
 FT CARBOHYD 648 648 POTENTIAL.
 SQ SEQUENCE 867 AA: 98399 MW: 7233587A CPG32;

Query Match 100.0%; Score 52; DB 1; Length 867;
 Best Local Similarity 100.0%; Pred. No. 0.0091;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TITFCASDAK 10
 |||||
 DB 49 TITFCASDAK 58

RESULT 10
 ENV_HY1JR STANDARD: PRT: 848 AA.
 ID ENV_HY1JR
 AC P20871;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ENVELOPE POLYPEPTIDE GP160 PRECURSOR [CONTAINS: EXTERIOR MEMBRANE
 DE GLYCOPROTEIN (GP120); TRANSMEMBRANE GLYCOPROTEIN (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (JRCSF isolate) (HIV-1).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 PA KOVANAOT S. CHEN I S Y.
 RL Submitted (DEC-1988) to the HIV data bank.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M38429; AAB03749.1;
 DR HIV: M38429; ENVSJCSF;
 DR PFAM: PF00517; GP120.1;
 DR PFAM: PF00517; GP41.1;
 KW Aids, Coat protein, Polypeptide, Glycoprotein, Transmembrane;
 KW Signal.
 FT SIGNAL 1 32
 FT CHAIN 33 503 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 504 848 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 53 73 BY SIMILARITY.
 FT DISULFID 118 203 BY SIMILARITY.
 FT DISULFID 125 194 BY SIMILARITY.
 FT DISULFID 130 154 BY SIMILARITY.
 FT DISULFID 216 245 BY SIMILARITY.

FT DISULFID 226 237 BY SIMILARITY.
 FT DISULFID 294 328 BY SIMILARITY.
 FT DISULFID 374 437 BY SIMILARITY.
 FT DISULFID 381 410 BY SIMILARITY.
 FT CARBOHYD 87 87 POTENTIAL.
 FT CARBOHYD 134 134 POTENTIAL.
 FT CARBOHYD 137 137 POTENTIAL.
 FT CARBOHYD 153 153 POTENTIAL.
 FT CARBOHYD 157 157 POTENTIAL.
 FT CARBOHYD 185 185 POTENTIAL.
 FT CARBOHYD 195 195 POTENTIAL.
 FT CARBOHYD 228 228 POTENTIAL.
 FT CARBOHYD 239 239 POTENTIAL.
 FT CARBOHYD 260 260 POTENTIAL.
 FT CARBOHYD 274 274 POTENTIAL.
 FT CARBOHYD 287 287 POTENTIAL.
 FT CARBOHYD 293 293 POTENTIAL.
 FT CARBOHYD 299 299 POTENTIAL.
 FT CARBOHYD 329 329 POTENTIAL.
 FT CARBOHYD 336 336 POTENTIAL.
 FT CARBOHYD 352 352 POTENTIAL.
 FT CARBOHYD 382 382 POTENTIAL.
 FT CARBOHYD 388 388 POTENTIAL.
 FT CARBOHYD 392 392 POTENTIAL.
 FT CARBOHYD 403 403 POTENTIAL.
 FT CARBOHYD 440 440 POTENTIAL.
 FT CARBOHYD 453 453 POTENTIAL.
 FT CARBOHYD 603 603 POTENTIAL.
 FT CARBOHYD 608 608 POTENTIAL.
 FT CARBOHYD 617 617 POTENTIAL.
 FT CARBOHYD 629 629 POTENTIAL.
 FT CARBOHYD 808 808 POTENTIAL.
 SQ SEQUENCE 848 AA: 96475 MW: A1EC8FA3 CPG32;

Query Match 100.0%; Score 52; DB 1; Length 848;
 Best Local Similarity 100.0%; Pred. No. 0.0089;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TITFCASDAK 10
 |||||
 DB 49 TITFCASDAK 58

RESULT 11
 ENV_HY1KB STANDARD: PRT: 861 AA.
 ID ENV_HY1KB
 AC P31819;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ENVELOPE POLYPEPTIDE GP160 PRECURSOR [CONTAINS: EXTERIOR MEMBRANE
 DE GLYCOPROTEIN (GP120); TRANSMEMBRANE GLYCOPROTEIN (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (KB-1 isolate) (HIV-1).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92351552.
 RA SHIMIZU H., HASEBE F., USHIOKE H., MURIKAWA S., USHIJIMA H.,
 RA KITAMURA T.;
 RT "Analysis of a human immunodeficiency virus type 1 isolate carrying a
 RT truncated transmembrane glycoprotein";
 RL Virology 189:534-546(1992).
 CC -1- MISCELLANEOUS: REENTRY THROUGH OF THE TERMINATOR TAG OCCURS BETWEEN
 CC THE CODONS FOR 729-ALA AND 730-ARG.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@slb.ch).

CC EMBL: D12582; BAAC214.1; ALT_SEU.
 DR PIR: A42995; VGLJRX.
 DR PIR: B42995; VGLJRX.
 DR PFAM: PF00517; GP41; 1.
 DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
 KW Signal.
 FT SIGNAL 1 35
 FT CHAIN 1 517
 FT CHAIN 518 861
 FT TRANSMEM 690 711
 FT TRANSMEM 711 79
 FT DISULFID 124 212
 FT DISULFID 131 203
 FT DISULFID 141 160
 FT DISULFID 225 254
 FT DISULFID 235 245
 FT DISULFID 303 337
 FT DISULFID 383 446
 FT DISULFID 490 419
 FT CARBOHYD 93 93
 FT CARBOHYD 141 141
 FT CARBOHYD 145 145
 FT CARBOHYD 146 146
 FT CARBOHYD 159 159
 FT CARBOHYD 163 163
 FT CARBOHYD 191 191
 FT CARBOHYD 192 192
 FT CARBOHYD 237 237
 FT CARBOHYD 241 241
 FT CARBOHYD 248 248
 FT CARBOHYD 259 259
 FT CARBOHYD 283 283
 FT CARBOHYD 296 296
 FT CARBOHYD 308 308
 FT CARBOHYD 348 348
 FT CARBOHYD 348 348
 FT CARBOHYD 361 361
 FT CARBOHYD 367 367
 FT CARBOHYD 397 397
 FT CARBOHYD 403 403
 FT CARBOHYD 408 408
 FT CARBOHYD 414 414
 FT CARBOHYD 449 449
 FT CARBOHYD 465 465
 FT CARBOHYD 468 468
 FT CARBOHYD 517 517
 FT CARBOHYD 622 622
 FT CARBOHYD 631 631
 FT CARBOHYD 643 643
 FT CARBOHYD 821 821
 SU SEQUENCE 861 AA: 98116 MW: 64D74FOF CRC32:

Query Match 100.0%; Score 52; DB 1; Length 861;
 Best Local Similarity 100.0%; Pred. No. 0; Inp91;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 1 TITFCASDAK 10
 DB 55 TITFCASDAK 64

RESULT 12
 ENV_HVILW STANDARD: PRT: 856 AA.
 AT 070625;
 DT 15-JUL-1998 (rel. 35, Created)
 DT 15-JUL-1998 (rel. 35, Last sequence update)
 DT 15-JUL-1999 (rel. 38, Last annotation update)
 DE ENVELOPE POLYPROTEIN GP160 PRECURSOR (CONTAINS: EXTERIOR MEMBRANE

DE GLYCOPROTEIN (GP120); TRANSMEMBRANE GLYCOPROTEIN (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (TM12.3 isolate) (HIV-1).
 CC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 CC [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE: 95127297.
 RA REITZ M.S., JR., HALL L., ROBERT-GUPOFF M., LAUTENBERGER J., HAHN B.M.,
 RA SHAW G.M., KONG L.I., WEISS S.H., WATERS D., GALLO R.C., BLATTNER W.J.
 RT "Viral variability and serum antibody response in a laboratory worker
 infected with HIV type 1 (HIV type 11B)."
 RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/infocentre>
 CC or send an email to license@slb.ch).

CC EMBL: U12055; AAA76690.1; -.
 DR PFAM: PF00516; GP120; 1.
 DR PFAM: PF00517; GP41; 1.
 DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
 KW Signal.
 FT SIGNAL 1 30
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 166 166
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CARBOHYD 356 356
 FT CARBOHYD 386 386
 FT CARBOHYD 392 392
 FT CARBOHYD 397 397
 FT CARBOHYD 406 406
 FT CARBOHYD 448 448
 FT CARBOHYD 463 463
 FT CARBOHYD 611 611
 FT CARBOHYD 616 616
 FT CARBOHYD 624 624
 FT CARBOHYD 637 637
 FT CARBOHYD 674 674
 FT CARBOHYD 750 750
 FT CARBOHYD 816 816
 SU SEQUENCE 856 AA: 96938 MW: C0AE181B CRC32:

Query Match 100.0%; Score 52; DB 1; Length 856;

Best Local Similarity: 100.0%; Pred. No. 0.009;
Matches: 10; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0

```
QY      1 TILFCASDAK 10
        |||||
Db      50 TILFCASDAK 59
```

RESULT 13
ENV HY1MA

ID	ENV_HVIMA	STANDARD;	PRT;	859 AA
AC	P04582			

DT	13-AUG-1987	(Rel. 05, Created)
DT	13-AUG-1987	(Rel. 05, Created)

DT 15-JUL-1999 (Rel. 38, Last annotation update
ENVELOPE POLYPEPTIDE 30160 ENVELOPE POLYPEPTIDE 30160

OF GLYCOTIDE (3P136), TRANSMEMBRANE GLYCOPROTEIN (3P41).

Human immunodeficiency virus type 1 (MAL isolate) (HIV-1)

RN [1]
PD SENTENCE FROM N A

BA	ALTON M	EATN-HOBSON C	WOUTCUTT D	COVITO P
PX	MERTINE:	86245056.		

Genetic variability of the AIDS virus: nucleotide sequence analysis of two isolates from African patients

Cell 46:63-74(1986)

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation in Grenoble.

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is not republished.

CC modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.ichsib.ch/announcements>.)

```
cc or send an email to license@isp-slb.cn).
cc -----
```

DR	EMBL:	X04415;	CAA28016.1;	-;
DB	EMBL:	A07115;	CA000623.1;	-;

DR HL V; K03456; ENVSMAT
DR PEAM: PF00516: GP120

DR PEAM; PE00517; GP41; 1.
KW AIDS: Coat protein: pol

ET SIGNAL 1 30

CHAIN	31	313	EXTERIOR MEMBRANE GLYCOPROTEIN.
ET CHAIN	514	859	TRANSMEMBRANE GLYCOPROTEIN

ET	118	210
DISULFID	33	73
E1	33	73

E1	DISULFID	133	201
E1	DISULFID	130	162

ET	DISULFID	233	244
ET	DISULFID	233	244

ET	DISULEID	380	445
ET	DISULEID	301	334

FT	CARBOHYD	87	87
FT	DISOLFD	387	418

FT	125	134
CARBOHYD	125	134
CHLORIDE	125	134

FT	146	146
CARBOHYD	146	146

FT	CARECHYD	193	193
----	----------	-----	-----

FT	239	239
CARBOHYD	239	239

FT	CARBOHYD	267	267

FT	CARBONHD	294	294
----	----------	-----	-----

FT	306	306
CARBOHYD	306	306

FT	CARBOHYD	357	357
----	----------	-----	-----

FT	CARBOHYD	388	388	POTENTIAL.
FT	CARBOHYD	394	394	POTENTIAL.
FT	CARBOHYD	406	406	POTENTIAL.
FT	CARBOHYD	448	448	POTENTIAL.
FT	CARBOHYD	460	460	POTENTIAL.
FT	CARBOHYD	467	467	POTENTIAL.
FT	CARBOHYD	613	613	POTENTIAL.
FT	CARBOHYD	618	618	POTENTIAL.
FT	CARBOHYD	627	627	POTENTIAL.
FT	CARBOHYD	639	639	POTENTIAL.
FT	CARBOHYD	819	819	POTENTIAL.
SQ	SEQUENCE	859 AA:	97100 MW:	1A2CDB71 C9C32:

Query Match	Score	EB	Length
Best Local Similarity	100.0%	113.0	0.003
Matches	10	0	0
Mismatches	0	0	0
Gaps	0	0	0

RESULT	14
ENV_HYMF	
ID	ENV_HYMF
	STANDARD:
	PRT:
	853 AA.
AC	P19551
DT	01-FEB-1991 (Rel. 17, Created)
DT	01-FEB-1991 (Rel. 17, last sequence update)
DT	15-JUL-1999 (Rel. 38, last annotation update)

DE GLYCOPROTEIN (GP120); TRANSMEMBRANE GLYCOPROTEIN (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (MFA isolate) (HIV-1).
PA Viruses; Retroviral viruses; Retroviridae; Lentiviruses.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90317877.
TX STEVENSON M., HAGGERTY S., IAMBONICA C., MANN A.M., MEIER C.,

RT "Colloids and characterization of human immunodeficiency virus type I
RT variants diminished in the ability to induce syncytium-independent
RT cytolysis." *J. Virol.* 64:3792-3803(1990)
CC
CC
CC This SWISS-PROT entry is copyright. It is provided through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. Where appropriate, the
CC

CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (see <http://www.int-sit.org/about-int-sit/>
CC or send an email to license@int-sit.ch).
DR EMBL_M33943; AAAA4850.1; -
DR HIV; M33943; ENVSTRA;
DR PRAM; PF00216; GP120; 1.
CC

NM	Aldo ₁ , Coat Protein; Glycoprotein; Transmembrane Signal.
KM	Signal.
FI	CHAIN 1 30
FI	CHAIN 31 509
FI	CHAIN 510 853
FI	DISULFID 54 BY SIMILARITY.
FI	DISULFID 74 BY SIMILARITY.
FI	DISULFID 119 203 BY SIMILARITY.
FI	DISULFID 126 194 BY SIMILARITY.
FI	DISULFID 126 194 BY SIMILARITY.

FT	DISULFID	216	245	BY SIMILARITY.
FT	DISULFID	226	237	BY SIMILARITY.
FT	DISULFID	294	339	BY SIMILARITY.
FT	DISULFID	376	443	BY SIMILARITY.
FT	DISULFID	383	416	BY SIMILARITY.
FT	CARBOHYD	88	88	POTENTIAL.
FT	CARBOHYD	136	136	POTENTIAL.

FT	CARBOHYD	141	141	POTENTIAL.
ET	CARBOHYD	156	156	POTENTIAL.
ET	CARBOHYD	160	160	POTENTIAL.
ET	CARBOHYD	186	186	POTENTIAL.
ET	CARBOHYD	195	195	POTENTIAL.
ET	CARBOHYD	242	242	POTENTIAL.
ET	CARBOHYD	249	239	POTENTIAL.
ET	CARBOHYD	260	260	POTENTIAL.
ET	CARBOHYD	274	274	POTENTIAL.
ET	CARBOHYD	297	297	POTENTIAL.
ET	CARBOHYD	293	293	POTENTIAL.
ET	CARBOHYD	299	299	POTENTIAL.
ET	CARBOHYD	340	340	POTENTIAL.
ET	CARBOHYD	354	354	POTENTIAL.
ET	CARBOHYD	384	384	POTENTIAL.
ET	CARBOHYD	390	390	POTENTIAL.
ET	CARBOHYD	395	395	POTENTIAL.
ET	CARBOHYD	404	404	POTENTIAL.
ET	CARBOHYD	446	446	POTENTIAL.
ET	CARBOHYD	461	461	POTENTIAL.
ET	CARBOHYD	609	609	POTENTIAL.
ET	CARBOHYD	614	614	POTENTIAL.
ET	CARBOHYD	623	623	POTENTIAL.
ET	CARBOHYD	635	635	POTENTIAL.
ET	CARBOHYD	672	672	POTENTIAL.
ET	CARBOHYD	748	748	POTENTIAL.
ET	CARBOHYD	814	814	POTENTIAL.
SO	SPECIMENF	H ₂ O AN:	964.12 MM:	RACGAAAF CFC12:

```
QY      1  TTI,FCASIDAK  10
        1 111111111
DB      50  TTI,FCASIDAK  59
```

RESULT	15
ENV_HYLMN	
AC	POS877:
ID	STANDARD:
DT	PT: 856 AA.
DT	01-NOV-1988 (Rel. 09, Created)
DT	01-NOV-1988 (Rel. 09, Last sequence update)
DT	15-JUL-1999 (Rel. 48, Last annotation update)
DE	ENVELOPE POLYPEPTIDE gp16 PREPROTEIN [CONTAINS: EXTERIOR MEMBRANE
DEF	GLYCOPROTEIN (GP120) : TRANSMEMBRANE GLYCOPROTEIN (GP41)].

-- MISCELLANEOUS: THE MN ISOLATE WAS TAKEN FROM A PEDIATRIC AIDS PATIENT IN 1984.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by, and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: M17449; AAA44857.1; -
DR PIR: A28922; VCLJMN.
DR HIV: M17449; ENVS MN.
DR PFAM: PF00516; GP120; 1.

[illegible]

```
QY      1  TTLEFCASDAK 10
Db      49  TTLEFCASDAK 58
```

Search completed: December 9, 1999, 22:12:24
Job time: 1526 sec

Wed Dec 15 08:56:44 1999

us-09-017-524-32.rsp

Page 11



GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 1999, 21 07:49 : Search time: 147.28 Seconds
(without alignments)
4159 Million cell updates/sec

Title: us-09-017-524-32
Perfect score: 52
Sequence: 1 TITLFCASDAK 10

Scoring table: RtoSum62

Searched: 199794 seqs, 61255205 residues

Database:

SPTEMBL.11:
1: sp.archaea:
2: sp.bacteria:
3: sp.fungi:
4: sp.human:
5: sp.invertebrate:
6: sp.mammal:
7: sp.mhc:
8: sp.organelle:
9: sp.phage:
10: sp.plant:
11: sp.rodent:
12: sp.virus:
13: sp.vertebrate:
14: sp.unclassified:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	52	100.0	795	2	Q46469	Q46469 escherichia
2	52	100.0	853	12	055774	055774 human immun
3	52	100.0	844	12	056561	056561 human immun
4	52	100.0	851	12	056562	056562 human immun
5	52	100.0	853	12	056563	056563 human immun
6	52	100.0	850	12	056564	056564 human immun
7	52	100.0	854	12	066566	066566 human immun
8	52	100.0	852	12	056567	056567 human immun
9	52	100.0	473	12	057037	057037 human immun
10	52	100.0	478	12	057038	057038 human immun
11	52	100.0	872	12	071280	071280 human immun
12	52	100.0	855	12	072612	072612 human immun
13	52	100.0	845	12	070679	070679 human immun
14	52	100.0	842	12	070895	070895 human immun
15	52	100.0	845	12	070902	070902 human immun
16	52	100.0	850	12	071250	071250 human immun
17	52	100.0	851	12	071270	071270 human immun
18	52	100.0	856	12	071275	071275 human immun
19	52	100.0	858	12	071974	071974 human immun
20	52	100.0	857	12	072605	072605 human immun
21	52	100.0	853	12	066108	066108 human immun
22	52	100.0	851	12	056110	056110 human immun
23	52	100.0	846	12	056111	056111 human immun
24	52	100.0	854	12	056112	056112 human immun
25	52	100.0	850	12	056113	056113 human immun
26	52	100.0	854	12	056319	056319 simian-huma
27	52	100.0	840	12	070694	070694 human immun
28	52	100.0	480	12	071152	071152 human immun
29	52	100.0	484	12	071156	071156 human immun

ALIGNMENTS

30	52	100.0	829	12	Q76123	Q76123 human immun
31	52	100.0	862	12	Q42069	Q42069 human immun
32	52	100.0	475	12	Q71144	Q71144 human immun
33	52	100.0	478	12	Q71146	Q71146 human immun
34	52	100.0	452	12	Q71148	Q71148 human immun
35	52	100.0	495	12	Q71150	Q71150 human immun
36	52	100.0	489	12	Q71154	Q71154 human immun
37	52	100.0	480	12	Q71159	Q71159 human immun
38	52	100.0	807	12	Q76122	Q76122 human immun
39	52	100.0	859	12	Q66225	Q66225 human immun
40	52	100.0	497	12	Q78694	Q78694 human immun
41	52	100.0	497	12	Q78595	Q78595 human immun
42	52	100.0	495	12	Q78597	Q78597 human immun
43	52	100.0	499	12	Q78598	Q78598 human immun
44	52	100.0	497	12	Q78600	Q78600 human immun
45	52	100.0	348	12	Q74999	Q74999 human immun

RESULT 1
Q46469 ID Q46469 PRELIMINARY: PRT: 795 AA.

AC Q46469: 01-NOV-1996 (TREMBL:01, Created)
DT 01-NOV-1996 (TREMBL:01, Last sequence update)
DT 01-MAY-1999 (TREMBL:10, Last annotation update)
DE HIV ENVELOP PROTEIN gp160
OS Escherichia coli.
OC Plasmid pSP16
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RA BUCKLER A.J.:
RL Submitted (MAY-1995) to the EMBL/JenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94184375.
RA CHURCH D.M., STOTLER C.J., PUTTER J.L., MURRELL J.P., TROTTIER J.A.,
RA BUCKLER A.J.:
RT "Isolation of genes from complex sources of mammalian genomic DNA
RT using exon amplification."
RL Nat. Genet. 6:98-105(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97171052.
RA NISSON P.E., ALDY A., WATKINS P.C.:
RT "Protocols for trapping internal and 3'-terminal exons."
RL PCR Methods Appl. 4:524-539(1994).
RN [4]
RP SEQUENCE FROM N.A.
RA HORTON L.:
RL Submitted (JAN-1995) to the EMBL/JenBank/DBJ databases.
PL EMBL; U19867; AAA64565.1.
DR EMBL; PF00516; GP150.1.
DR EMBL; PF00517; GP150.1.
DR EMBL; PF00424; REV.1.
KW Plasmid.
SQ SEQUENCE 795 AA. 89945 MW. 2463534.30000:

Query Match 100.0%: Score 52, E-2, Length 795,
Best Local Similarity 100.0%: Pred. No. 0.0082,
Matches 10: Conservative 0, Mismatches 0, Indels 0, Gaps 0;

QY 1 TITLFCASDAK 10
|||||
Db 72 TITLFCASDAK 81

RESULT 2

055774 PRELIMINARY: PRT: 853 AA.
 ID 055774
 AC 055774:
 RN [1]
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)
 DE ENVELOPE GLYCOPROTEIN.
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses: Retroid viruses; Retroviridae; Lentivirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Q23-17:
 RA POSS M., DYERBAUGH J.:
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF004485; AAC07527.1;
 DR PFAM: PF00516; GP120.1.
 DR PFAM: PF00517; GP41.1.
 DR ENVELOPE PROTEIN.
 KW SEQUENCE 853 AA: 96778 MW: 99324480 CRC32:
 SO

Query Match 100.0% Score 52: DB 12: Length 853:
 Best Local Similarity 100.0% Pred. No. 0.0087:
 Matches 10: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 TTLCASDAK 10
 DB 49 TTLCASDAK 58

RESULT 4
 ID 056561 PRELIMINARY: PRT: 844 AA.
 AC 056561:
 RN [1]
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)
 DE ENVELOPE GLYCOPROTEIN.
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses: Retroid viruses; Retroviridae; Lentivirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PH1153:
 RA ATAMAN-ONAL Y., GIRAUD A., COIFFIER C., BABIC-EPCEG A., BIRON F.,
 RA VERRIER B.:
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF041124; AAC02517.1;
 DR PFAM: PF00516; GP120.1.
 DR PFAM: PF00517; GP41.1.
 DR ENVELOPE PROTEIN.
 KW SEQUENCE 844 AA: 95831 MW: 74605490 CRC32:
 SO

Query Match 100.0% Score 52: DB 12: Length 844:
 Best Local Similarity 100.0% Pred. No. 0.0085:
 Matches 10: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 TTLCASDAK 10
 DB 49 TTLCASDAK 58

RESULT 4
 ID 056562 PRELIMINARY: PRT: 851 AA.
 AC 056562:
 RN [1]
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)
 DE ENVELOPE GLYCOPROTEIN.
 GN ENV.

OS Human immunodeficiency virus type 1.
 OC Viruses: Retroid viruses; Retroviridae; Lentivirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PH1153:
 RA ATAMAN-ONAL Y., GIRAUD A., COIFFIER C., BABIC-EPCEG A., BIRON F.,
 RA VERRIER B.:
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF041128; AAC02519.1;
 DR PFAM: PF00516; GP120.1.
 DR PFAM: PF00517; GP41.1.
 DR ENVELOPE PROTEIN.
 KW SEQUENCE 851 AA: 96811 MW: 594D3997 CRC32:
 SO

Query Match 100.0% Score 52: DB 12: Length 851:
 Best Local Similarity 100.0% Pred. No. 0.0087:
 Matches 10: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 TTLCASDAK 10
 DB 49 TTLCASDAK 58

RESULT 5
 ID 056563 PRELIMINARY: PRT: 853 AA.
 AC 056563:
 RN [1]
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)
 DE ENVELOPE GLYCOPROTEIN.
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses: Retroid viruses; Retroviridae; Lentivirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PH1153:
 RA ATAMAN-ONAL Y., GIRAUD A., COIFFIER C., BABIC-EPCEG A., BIRON F.,
 RA VERRIER B.:
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF041129; AAC02520.1;
 DR PFAM: PF00516; GP120.1.
 DR PFAM: PF00517; GP41.1.
 DR ENVELOPE PROTEIN.
 KW SEQUENCE 853 AA: 96799 MW: 4105BD18 CRC32:
 SO

Query Match 100.0% Score 52: DB 12: Length 853:
 Best Local Similarity 100.0% Pred. No. 0.0087:
 Matches 10: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 TTLCASDAK 10
 DB 50 TTLCASDAK 59

RESULT 6
 ID 056564 PRELIMINARY: PRT: 850 AA.
 AC 056564:
 RN [1]
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)
 DE ENVELOPE GLYCOPROTEIN.
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses: Retroid viruses; Retroviridae; Lentivirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PH1155:
 RA ATAMAN-ONAL Y., GIRAUD A., COIFFIER C., BABIC-EPCEG A., BIRON F.,
 RA VERRIER B.:

Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF041130; AAC02521.1;
 DR PFAM: PF00516; GP120; 1.
 DR EMBL: AF00517; GP41; 1.
 KW Envelope protein.
 SQ SEQUENCE 850 AA; 96294 MW; E95E960C CRC32;

Query Match 100.0%; Score 52; DB 12; Length 850;
 Best Local Similarity 100.0%; Pred. No. 0.0087;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TITFCASDAK 10
 |||||||
 DB 49 TITFCASDAK 58

RESULT 7
 ID 056566 PRELIMINARY; PRI: 854 AA.
 AC 056566;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE ENVELOPE GLYCOPROTEIN.
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PIH309;
 RA ATAMAN-ONAL Y., GIFFAUD A., COIFFIER C., PARIC-EPCEG A., BIRON F.,
 RA VERPIER B.;
 RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF041132; AAC02523.1;
 DR PFAM: PF00516; GP120; 1.
 DR PFAM: PF00517; GP41; 1.
 KW Envelope protein.
 SQ SEQUENCE 854 AA; 97062 MW; 6570F5E9 CRC32;

Query Match 100.0%; Score 52; DB 12; Length 854;
 Best Local Similarity 100.0%; Pred. No. 0.0087;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TITFCASDAK 10
 |||||||
 DB 50 TITFCASDAK 59

RESULT 8
 ID 056567 PRELIMINARY; PRI: 852 AA.
 AC 056567;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE ENVELOPE GLYCOPROTEIN.
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PIH374;
 RA ATAMAN-ONAL Y., GIFFAUD A., COIFFIER C., PARIC-EPCEG A., BIRON F.,
 RA VERPIER B.;
 RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF041133; AAC02524.1;
 DR PFAM: PF00516; GP120; 1.
 DR PFAM: PF00517; GP41; 1.
 KW Envelope protein.
 SQ SEQUENCE 852 AA; 96769 MW; C13815F CRC32;

Query Match 100.0%; Score 52; DB 12; Length 852;
 Best Local Similarity 100.0%; Pred. No. 0.0087;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TITFCASDAK 10
 |||||||
 DB 49 TITFCASDAK 58

RESULT 9
 ID 057037 PRELIMINARY; PRI: 473 AA.
 AC 057037;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE ENVELOPE GLYCOPROTEIN (FRAGMENT)
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=028;
 RA CORNELISSEN M., GOUDSMIT J.;
 RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL: U16217; AAC01499.1;
 DR PFAM: PF00516; GP120; 1.
 KW Envelope protein.
 FT NON_TER 1
 FT NON_TER 473
 FT NON_TER 473
 SQ SEQUENCE 473 AA; 53362 MW; 744B8B8C CRC32;

Query Match 100.0%; Score 52; DB 12; Length 473;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TITFCASDAK 10
 |||||||
 DB 19 TITFCASDAK 28

RESULT 10
 ID 057038 PRELIMINARY; PRI: 478 AA.
 AC 057038;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE ENVELOPE GLYCOPROTEIN (FRAGMENT)
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=028;
 RA CORNELISSEN M., GOUDSMIT J.;
 RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL: U16218; AAC01500.1;
 DR PFAM: PF00516; GP120; 1.
 KW Envelope protein.
 FT NON_TER 1
 FT NON_TER 478

FT N/NTER 478 478
 SO SEQUENCE 478 AA: 54180 MW: 20573800 CRC32:

Query Match 100.0%; Score 52; DB 12; Length 478;
 Best Local Similarity 100.0%; Pred. No. 0.0052;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 1 TLFCASDAK 10
 |||||
 DB 19 TLFCASDAK 28

RESULT 11

ID 071280 PRELIMINARY: PRT: 872 AA.

AC 071280:

DT 01-AUG-1998 (TREMBLrel. 07, Created)

DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)

DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)

DE ENVELOPE GLYCOPROTEIN.

GN ENV.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-MBCD36:

RA OELRICHS R.H., MCPHER D.A., DEACON N.J.;

RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF042105; AAC0239.1; -

DR PFAM: PF00516; GP120.1;

DR PFAM: PF00517; GP41.1;

KW Envelope protein.

SO SEQUENCE 872 AA: 90254 MW: A815FC9 CPC32:

Query Match 100.0%; Score 52; DB 12; Length 872;
 Best Local Similarity 100.0%; Pred. No. 0.0089;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 1 TLFCASDAK 10
 |||||
 DB 50 TLFCASDAK 59

RESULT 12

ID 072619 PRELIMINARY: PRT: 855 AA.

AC 072619:

DT 01-AUG-1998 (TREMBLrel. 07, Created)

DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)

DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)

DE ENVELOPE GLYCOPROTEIN.

GN ENV.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-DJ264:

RA CARR J.K., SALMINEN M.O., ALBERT J., SANDERS-BUELL E., GOTTE D.,

RL Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF064224; AAC29084.1; -

DR PFAM: PF00516; GP120.1;

DR PFAM: PF00517; GP41.1;

KW Envelope protein.

SO SEQUENCE 855 AA: 97341 MW: 4029816 CPC32:

Query Match 100.0%; Score 52; DB 12; Length 855;
 Best Local Similarity 100.0%; Pred. No. 0.0087;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 1 TLFCASDAK 10
 |||||
 DB 48 TLFCASDAK 57

RESULT 13

ID 070679 PRELIMINARY: PRT: 845 AA.

AC 070679:

DT 01-AUG-1998 (TREMBLrel. 07, Created)

DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)

DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)

DE ENVELOPE GLYCOPROTEIN.

GN ENV.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-RL42:

RA GRAF M., SHAO Y., ZHAO Q., SEIDL T., KESTLER J., WOLF H., WAGNER R.;

RL Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.

DR EMBL: U71182; AAC05236.1; -

DR PFAM: PF00516; GP120.1;

DR PFAM: PF00517; GP41.1;

KW Envelope protein.

SO SEQUENCE 845 AA: 95963 MW: 2888E83 CRC32:

Query Match 100.0%; Score 52; DB 12; Length 845;
 Best Local Similarity 100.0%; Pred. No. 0.0086;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 1 TLFCASDAK 10
 |||||
 DB 49 TLFCASDAK 58

RESULT 14

ID 070895 PRELIMINARY: PRT: 842 AA.

AC 070895:

DT 01-AUG-1998 (TREMBLrel. 07, Created)

DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)

DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)

DE ENVELOPE GP160 PRECURSOR.

GN ENV.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-93BR029:

RA GAO F.;

RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF005495; AAC03179.1; -

DR PFAM: PF00516; GP120.1;

DR PFAM: PF00517; GP41.1;

SO SEQUENCE 842 AA: 94810 MW: 6594E468 CPC32:

Query Match 100.0%; Score 52; DB 12; Length 842;
 Best Local Similarity 100.0%; Pred. No. 0.0086;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TILFCASDAK 10
 |||||
 Db 48 TILFCASDAK 57

RESULT 15

070902
 ID 070902; PRELIMINARY; PRT; 845 AA.
 AC 070902;
 DI 01-AUG-1998 (TREMblrel. 07, Created)
 DI 01-AUG-1998 (TREMblrel. 07, last sequence update)
 DI 01-MAY-1999 (TREMblrel. 10, last annotation update)
 DE ENVELOPE POLYPROTEIN.
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=90CP056 (AKA U4056);
 RA GAO F.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF003496; AAD03181.1;
 DR PFAM; PF00516; GP120; 1.
 DR PFAM; PF00517; GP41; 1.
 KW Polyprotein; Envelope protein.
 SQ SPODNCE 845 AA; 95827 MW, 6380AF39 CRC32;

Query Match 100.0%; Score 52; DB 12; Length 845;
 Best Local Similarity 100.0%; Pred. No. 0.0086;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TILFCASDAK 10
 |||||
 Db 46 TILFCASDAK 55

Search Completed: December 9, 1999, 21:07:49
 Job time: 2306 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 1999, 20:57:29 ; Search time 137.64 Seconds
(without alignments)
1.549 Million cell updates/sec

Title: US-09-017-524-33
Perfect score: 42
Sequence: 1 RIVELLGRP 9

Scoring table: ELOSUM62

Searched: 188963 seqs, 23686106 residues

Database: A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	395	1 P91956	HIV p41 gene with
2	42	100.0	709	1 P95781	Fusion protein of
3	42	100.0	704	1 P95096	PSD302, pep HIV-1 g
4	42	100.0	412	1 P05095	Synthetic HIV-1 tr
5	42	100.0	32	1 P83470	Peptide component
6	42	100.0	36	1 P80666	HIV protein HT7, p
7	42	100.0	851	1 P80967	HIV protein HT6, p
8	42	100.0	1458	1 P80804	Sequence of HIV-1
9	42	100.0	853	1 P81858	Sequence encoded b
10	42	100.0	850	1 P81855	Sequence encoded b
11	42	100.0	845	1 P08406	Sequence deduced f
12	42	100.0	858	1 P60063	HIV virus env gene
13	42	100.0	865	1 P70175	Sequence encoded b
14	42	100.0	901	1 P70665	Sequence encoded b
15	42	100.0	857	1 P12261	HIV-1 strain GYI E
16	42	100.0	856	1 P61514	Sequence of envlo
17	42	100.0	858	1 P60432	Sequence of LAV vi
18	42	100.0	863	1 P60349	HIV-III virus (HI
19	42	100.0	856	1 P60131	Sequence of the AI
20	42	100.0	857	1 R14904	HIV-1(MN-STI) env
21	42	100.0	856	1 P25440	Modified HIV env g
22	42	100.0	863	1 P28955	Non-cleavable, sol
23	42	100.0	34	1 P33839	Cluster peptide PC
24	42	100.0	863	1 P43869	HIV-III ENV-LOR g
25	42	100.0	863	1 R43873	HIV-III ENV-LOR g
26	42	100.0	856	1 P41028	Selectively deglyc
27	42	100.0	856	1 P41026	Selectively deglyc
28	42	100.0	856	1 P41027	Selectively deglyc
29	42	100.0	856	1 R41028	Selectively deglyc
30	42	100.0	856	1 R41029	Selectively deglyc
31	42	100.0	856	1 R41030	Selectively deglyc
32	42	100.0	856	1 R41031	Selectively deglyc
33	42	100.0	856	1 R41032	Selectively deglyc
34	42	100.0	856	1 P45445	Sequence encoded b
35	42	100.0	856	1 P52687	gp41 HIV envelope
36	42	100.0	420	1 P53785	Translation of HIV
37	42	100.0	857	1 P67725	gp120 from the HIV
38	42	100.0	856	1 R67726	gp120 from the HIV
39	42	100.0	856	1 R67724	gp120 from the HIV
40	42	100.0	856	1 R73909	HIV-1 envelope pol
41	42	100.0	856	1 R79173	HIV virus-1 isolat
42	42	100.0	856	1 W79421	gp1 peptide 57, D
43	42	100.0	856	1 W1579	Human Immunodefici

ALIGNMENTS

44	42	100.0	857	1 W1580	Human Immunodefici
45	42	100.0	859	1 W72997	HIV isolate LAV.MA

RESULT 1	
ID P91956	P91956 standard; protein; 395 AA.
AC P91956;	
DT 10-MAR-1993 (revised)	
DT 13-FEB-1990 (first entry)	
DE HIV p41 gene with p120 carboxy terminus	
KW Fusion protein; CDS gene; Human immunodeficiency virus.	
FT key	Location/Qualifiers
FT misc-feature	5..43
FT	/note="carboxy terminus of p120"
FT misc-feature	194..231
FT	/note="hydrophobic region of p41"
FT misc-difference 74	
FT	/label= stop-codon
FT misc-difference 76	
FT	/label= stop-codon
FT misc-difference 107	
FT	/label= stop-codon
PN EP-331961-A.	
PD 13-SEP-1989.	
PF 20-FEB-1989; 102928.	
PR 11-MAR-1988; US-276263.	
PA (ABBO) Abbott Laboratories.	
PI Boiling TJ, Mandeckl W;	
DR WPI: 89-265215/37.	
N-PSDB: N90797.	
PT Expressing high levels of fusion protein in microbe . by transforming	
PT microbe with DNA vector having control region and part of CDS protein	
PT and expressing protein	
PS Disclosure; fig. 9; 24pp; English.	
CC The fusion gene of this is an expression vector entity, a lac	
CC control promoter and all/part of CDS (CDS: CDS: 336300-mannosyltransferase	
CC cytidyl transferase or CMP-KDO synthetase) protein.	
SC Sequence 395 AA;	

Query Match	100.0%	Score 42;	DB 1;	Length 395;
Best Local Similarity	100.0%	Pred. No. 0.33;		
Matches 9;	Conservative	0	Mismatches 0;	Gaps 0;

QY 1 RIVELLGRP 9	
Db 319 RIVELLGRP 327	

RESULT 2	
ID P95781	P95781 standard; protein; 709 AA.
AC P95781;	
DT 22-AUG-1990 (first entry)	
DE Fusion protein of HIV gag-env.	
KW HIV; fusion protein; gag; env; AIDS; vaccine; ss.	
OS Human T-cell lymphotropic virus III.	
PN J01179687-A.	
PD 17-JUL-1989.	
PF 30-DEC-1987; 336292.	
PR 30-DEC-1987; JP-336292.	
PA (KAGA-) Kaogaku Oyobi Kessei.	
DR WPI: 89-245894/34.	
N-PSDB: N90531.	
PT Fusion protein of HIV gag-env - effective for an AIDS	
PT antibody assay reagent or vaccine.	
PS Claim; 12pp; Japanese.	
CC Fusion protein gene comprises the C-94nt gag-env fragment from BgIII site	
CC to BgIII site of HIV-III gag gene, and 1.2kbp gene fragment from BgIII	

CC site to Xho I site of HIV-1 env gene. The HIV gag-env protein has the
 CC antigenicity of the gag and env proteins. It contains all the peptide
 CC domains of gag-p24 and env-gp41 that are esp. important antigens
 CC in a vaccine. The eukaryotic expression cell is pref. Saccharomyces
 CC cerevisiae.
 SQ Sequence 709 AA:

Query Match 100.0% Score 42; DB 1; Length 709;
 Best Local Similarity 100.0%; Pred. No. 0.61;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIVELLGRR 9
 |||||
 DB 643 RIVELLGRR 641

RESULT 4
 R05096
 ID R05096 standard; protein: 704 AA.
 AC R05096;
 DT 05-OCT-1990 (first entry)
 DE PSD302 PEP HIV-1 gag/env fusion protein
 KW recombinant HIV antigens; HIV-1; HIV-2; transmembrane glycoprotein;
 KM ss: HIV-1 gag/env fusion protein.
 OS synthetic.
 FH Key Location/Qualifiers
 FT Protein 13..299
 FT Protein /label=HIV-1 gag sequence
 FT Protein 405..695
 FT Protein /label=HIV-1 env sequence
 PN EP-370458-A.
 PD 40-MAY-1990.
 PE 21-NOV-1989; 121513.
 PR 23-NOV-1988; US-275309.
 PA (ABR) Abbott Laboratories.
 PI Devare SG, Suresh DM, Casey JM;
 DT WPI: 90-165199/22.
 FT Synthetic DNA derived recombinant HIV antigens -
 PT useful for detection of exposure to HIV as well as
 PT discrimination of HIV-1 and HIV-2
 PS Disclosure: P: English
 CC The fusion gene which codes for this protein (on plasmid PSD302)
 CC consists of the carboxyl-terminus of HIV-1 gp120 fused to the
 CC full-length HIV-1 gp41.
 CC See also 004470, 004471 and R05099.
 SQ Sequence 704 AA:

Query Match 100.0% Score 42; DB 1; Length 704;
 Best Local Similarity 100.0%; Pred. No. 0.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIVELLGRR 9
 |||||
 DB 619 RIVELLGRR 627

RESULT 4
 R05095
 ID R05095 standard; protein: 412 AA.
 AC R05095;
 DT 05-OCT-1990 (first entry)
 DE Synthetic HIV-1 transmembrane glycoprotein
 KW recombinant HIV antigens; HIV-1; HIV-2; transmembrane glycoprotein;
 OS synthetic.
 PN EP-370458-A.
 PD 40-MAY-1990.
 PE 21-NOV-1989; 121513.
 PR 23-NOV-1988; US-275309.
 PA (ABR) Abbott Laboratories.
 PI Devare SG, Suresh DM, Casey JM;

DR WPI: 90-165199/22.
 DR N-PSDB: 004470.
 PT Synthetic DNA derived recombinant HIV antigens -
 PT useful for detection of exposure to HIV as well as
 PT discrimination of HIV-1 and HIV-2
 PS Claim 1; Page 11; 33pp; English.
 CC This recombinant protein can be used to detect HIV antibodies. In
 CC conjunction with a recombinant HIV-2 antigen it can discriminate
 CC exposure to HIV-1 from HIV-2. Both recombinant antigens are
 CC expressed at high levels in E. coli.
 CC See also 004471, R05096 and R05099.
 SQ Sequence 412 AA:

Query Match 100.0% Score 42; DB 1; Length 412;
 Best Local Similarity 100.0%; Pred. No. 0.35;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIVELLGRR 9
 |||||
 DB 320 RIVELLGRR 328

RESULT 5
 P82470
 ID P82470 standard; protein: 32 AA.
 AC P82470;
 DT 12-NOV-1990 (first entry)
 DE Peptide component of AIDS vaccine.
 KM AIDS vaccine; T-cells.
 OS synthetic.
 PN EP-273716-A.
 PD 06-JUL-1988.
 PE 23-DEC-1987; 311391.
 PR 30-DEC-1986; US-947935.
 PR 12-FEB-1987; US-014430.
 PA (USDC) US Sec of Commerce.
 PA (USSH) US Dept Health and Human.
 PI Delisi C, Margalit H, Cornette JL, Ouyang CS;
 DR WPI: 88-184640/27.
 FT Synthetic peptide(s) as vaccines for AIDS - selected from peptide
 PT regions which can fold as a maximally amphipathic helix recognised
 PT by T cells.
 PS Claim 9; Page 10; 16pp; English.
 CC This peptide is a component of an AIDS vaccine. It can fold as a
 CC maximally amphipathic helix and is recognised by T-cells immune to
 CC the AIDS virus envelope protein.
 CC See also P82462-69 and P82471-79.
 SQ Sequence 32 AA:

Query Match 100.0% Score 42; DB 1; Length 32;
 Best Local Similarity 100.0%; Pred. No. 0.024;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIVELLGRR 9
 |||||
 DB 4 RIVELLGRR 12

RESULT 6
 P80966
 ID P80966 standard; protein: 866 AA.
 AC P80966;
 DT 15-NOV-1990 (first entry)
 DE HIV protein HT7.
 KM HIV: HT7; gp160; envelope protein; RF; AIDS.
 OS Human immunodeficiency virus variant RF.
 PN EP-272858-A.
 PD 29-JUN-1988.
 PE 14-DEC-1987; 310967.
 PR 15-DEC-1986; US-941111.
 PR 31-AUG-1987; US091481.

PA (RPD-) PepLigen Corp.
 PI Rusche J, Lynn D, Carson H, Putney S, Jellis CL;
 DR WPI: 88-176944/26.
 DK N-PSDB; N80948.
 PT Prodn. of recombinant HIV envelope proteins in insect cells -
 PI useful as vaccine against AIDS and for diagnosis and therapy.
 PS Disclosure; App; English.
 CC GP 160 from HIV PF was cloned by ligating a fragment encoding the
 CC N-terminal AAs up to the DraIII site and C-terminal AAs beyond the
 CC Aval site (from P6160), and a second fragment encoding the major
 CC portion of the gene from the HIV PF variant. The resulting chim-
 CC eric gene, PH17, was subcloned into pAC510 to construct PH17. The
 CC protein has novel immunological and antigenic characteristics and
 CC may be used as vaccine or for diagnosis and therapy of AIDS.
 CC See also P80967
 SQ Sequence 856 AA;

Query Match 100.0%, Score 42, DB 1, Length 866;
 Best Local Similarity 100.0%; Pred. No. 0.75;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIVELLGRP 9
 DB 790 RIVELLGRP 798
 SQ

RESULT 7
 ID P80967 standard; protein, 851 AA.
 AC P80967;
 DT 15-NOV-1990 (first entry)
 DE HIV protein HT6.
 KM HIV; HT6; gp160; envelope protein; RF; AIDS.
 OS Human immunodeficiency virus variant RF.
 PN EP-272858-A.
 PD 29-JUN-1988.
 PF 14-DEC-1987; 310967.
 PR 15-DEC-1986; US-941111.
 PR 31-AUG-1987; US091481.
 PA (REPL-) Repligen Corp.
 PI Rusche J, Lynn D, Carson H, Putney S, Jellis CL;
 DR WPI: 88-176944/26.
 DK N-PSDB; N80948.
 PT Prodn. of recombinant HIV envelope proteins in insect cells -
 PT useful as vaccine against AIDS and for diagnosis and therapy.
 PS Disclosure; App; English.
 CC The sequence is the result of cloning a hybrid envelope gene from
 CC HIV variants BH10 and RF. A central portion of the RF gene was
 CC used, the rest being from the distantly related variant BH10. The
 CC resulting clone, PACH16 produces a hybrid gp 160 envelope protein
 CC with novel immunological and antigenic characteristics. It may
 CC be used to as a vaccine and for diagnosis and therapy of AIDS.
 CC See also P80966.
 SQ Sequence 851 AA;

Query Match 100.0%, Score 42, DB 1, Length 851;
 Best Local Similarity 100.0%; Pred. No. 0.74;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIVELLGRP 9
 DB 775 RIVELLGRP 783
 SQ

RESULT 8
 ID P80804 standard; protein, 1458 AA.
 AC P80804;
 DT 15-NOV-1990 (first entry)
 DE Sequence of HIV-1 BRU isolate CNCM number I-232 in
 DE clone HIV-1 P

KM AIDS; immunogen; antigen; vaccine; diagnostic.
 OS Human immunodeficiency virus 1 BRU.
 PN W08805440-A.
 PD 28-JUL-1988.
 PF 15-JAN-1988; F00025.
 PR 15-APR-1987; FR-005398,
 PA (INSP) Inst Pasteur.
 PI Alizon M, Montagnier L, Charette D, Charrel F, Sauter P, Guzman M,
 PI Riollan P, Charabarti L, Desrochers R;
 DR WPI: 88-220290/31.
 PT New peptides(s) with immunological properties of HIV-2 envelope protein -
 PT having the structure of simian immune deficiency virus proteins,
 PT useful in diagnosis and of vaccine components
 PS Disclosure, Fig 2, 86pp, French.
 CC The SQ of env protein of HIV-2 P0D (P80803) was compared with the SQ of
 CC env protein of HIV-1 BRU (P80804) in Fig 2. New peptides which have
 CC immunological properties in common with those of the peptide skeleton of
 CC the envelope protein of HIV-2 and also have a peptide structure in common
 CC with that of HIV-1 glycoprotein are claimed. Antigenic and immunogenic
 CC conjugates conig. the peptides and a kit to detect HIV-2 in biological
 CC fluids are new. The peptides are useful for in vitro diagnosis of HIV-2
 CC infection and some of them can be used as components of immunogens and
 CC vaccines against HIV. Antibodies raised against them can be used for
 CC treatment of AIDS.
 SQ Sequence 1458 AA;

Query Match 100.0%, Score 42, DB 1, Length 1458,
 Best Local Similarity 100.0%; Pred. No. 1.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIVELLGRP 9
 DB 1382 RIVELLGRP 1390
 SQ

RESULT 9
 ID P81858 standard; protein, 853 AA.
 AC P81858;
 DT 16-DEC-1990 (first entry)
 DE Sequence encoded by LAV EL 1 ENV gene
 KM HIV, HTLV III; AIDS, diagnosis, vaccine; probe, hybridisation.
 OS Lymphadenopathy associated virus EL 1.
 PN W08707906-A.
 PD 30-DEC-1987.
 PF 22-JUN-1987; E00326.
 PF 23-JUN-1986; EP-401380.
 PA (INSP) Inst Pasteur.
 PI Alizon M, Sonigo P, Main-Hodson S, Montagnier L;
 DR WPI: 88-014396/02.
 DK N-PSDB; N80436.
 PT New variants of lymphadenopathy associated virus (LAV) -
 PT used for prodn. of DNA, antigens and antibodies used in
 PT diagnosis of AIDS and pre-AIDS
 PS Claim 8; Fig 7A-7J; 72pp; English.
 CC LAV EL 1 (N80436) and LAV MA 1 (N80437) were isolated from the peripheral
 CC blood lymphocytes of patients. Different AIDS virus isolates concerned
 CC are designated by 3 letters of the patients name. Stable probes including
 CC the DNA sequences or DNA proviruses in eg. biological samples. The proteins
 CC or peptides can be used for detection of antibodies induced in vivo and
 CC present in biological fluids. The DNA can also be used for the expression
 CC of LAV viral antigens for the prodn. of a vaccine against LAV. The
 CC polypeptides can also be used for the prodn. of antibodies for the
 CC detection of proteins related to the LAV viruses, partic. for diagnosis
 CC of AIDS or pre-AIDS.
 SQ Sequence 853 AA;

Query Match 100.0%, Score 42, DB 1, Length 853;
 Best Local Similarity 100.0%; Pred. No. 0.74;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIVELLGRR 9
 |||||
 DB 777 RIVELLGRR 785

RESULT 10

PIR865
 ID P81865 standard; protein: 859 AA.
 AC P81865;
 DT 16-DEC-1990 (first entry)
 DE Sequence encoded by LAV MA L ENV gene
 KW HIV; HTLV III; AIDS; diagnosis; vaccine; probe; hybridisation.
 OS Lymphadenopathy associated virus MA L.
 PN W08707906-A.
 PD 30-DEC-1987.
 PE 22-JUN-1987; E00326.
 PK 21-JUN-1986; EP-401380.
 PA (INSEF) Inst Pasteur.
 PI Allison M, Sonigo P, Wain-Hobson S, Montanier L;
 DB WPI: 88-014166/02.
 DR N-PSDB: N80437.
 PT New variants of lymphadenopathy associated virus (LAV) -
 PI used for prodn. of DNA, antigens and antibodies used in
 PI diagnosis of AIDS and pre-AIDS.
 CC LAV Fl. 1 (N80436) and LAV MA L (N80437) were isolated from the peripheral
 CC blood lymphocytes of patients. Different AIDS virus isolates concerned
 CC are designated by 3 letters of the patients name. Stable probes including
 CC the DNA sequences can be used for detection of the new LAV viruses or
 CC related viruses or DNA proviruses in eg. biological samples. The proteins
 CC or peptides can be used for detection of antibodies induced in vivo and
 CC present in biological fluids. The DNA can also be used for the expression
 CC of LAV viral antigens for the prodn. of a vaccine against LAV. The
 CC polypeptides can also be used for the prodn. of antibodies for the
 CC detection of proteins related to the LAV viruses, partic. for diagnosis
 CC of AIDS or pre-AIDS.
 SO Sequence 859 AA;

Query Match 100.0%; Score 42; DB 1; Length 859;
 Best Local Similarity 100.0%; Prod No. 0.74;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 11

K08406
 ID R08406 standard; protein: 846 AA.
 AC R08406;
 DT 27-FEB-1991 (first entry)
 DE Sequence deduced from env gene of HIV 1-NDK.
 KW Human immunodeficiency virus; AIDS.
 OS HIV 1-NDK.
 PN W09013630-A.
 PD 15-NOV-1990.
 PE 02-MAY-1990; E00312.
 PR 03-MAY-1989; FR-005914.
 PA (INSEF) INSEF INST NAT SANTE.
 PI Barre-Sinoussi F, Chermann JC, Devaux C, Rey F, Sire J;
 DB WPI: 90-361470/48.
 DR N-PSDB: 006635.
 PT New HIV-NDK retrovirus and protein component - used in vaccines
 PT against immuno-deficiency disorders and in raising Mabs for
 PI retro-virus detection in vivo.
 PS Disclosure: Fig 2; 37pp; French.
 CC The HIV NDK virus was isolated from peripheral blood lymphocytes of
 CC an AIDS patient. A genomic library was prepd. from DNA extracted
 CC from CEM cells infected with the virus. The library was screened

CC with a PB1 probe corresp. to a fragment from HIV 1. The virus is
 CC more cytopathic than other strains and is not inhibited by OKT4A.
 CC It has been deposited as CCMC I-857. The sequence can be used to
 CC express proteins useful for diagnosing the presence of NDK and
 CC related viruses and in vaccines against immunodeficiency diseases.
 CC See also R08040 and R08401-07.
 SO Sequence 846 AA;

Query Match 100.0%; Score 42; DB 1; Length 846;
 Best Local Similarity 100.0%; Prod. No. 0.73;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIVELLGRR 9
 |||||
 DB 770 RIVELLGRR 778

RESULT 12

P60063
 ID P60063 standard; Protein: 868 AA.
 AC P60063;
 DT 11-JUN-1991 (first entry)
 DE HIV virus env gene product.
 KW HIV virus; vaccine; AIDS.
 PN EP-201716-A.
 PD 20-NOV-1986.
 PE 07-APR-1986; 104741.
 PR 08-APR-1985; US-721237.
 PR 29-APR-1985; US-728052.
 PR 10-JUL-1985; US-753769.
 PR 07-NOV-1985; US-783299.
 PA (GENE-) GENETIC SYST CORP.
 PI Watanabe SM, Cosand W, Mcardle S, Travis BM;
 DR WPI: 86-306545/47.
 DR N-PSDB: N60077.
 PT New DNA sequence comprising part of env. region of LAV genome -
 PT and corresp. transformational protein prod., useful in vaccine
 PT against AIDS.
 PS Disclosure: Fig 8; 27pp; English.
 CC The protein is immunologically reactive with antibodies against
 CC HIV virus, and can be used in a vaccine composition for
 CC prevention against AIDS.
 SO Sequence 868 AA;

Query Match 100.0%; Score 42; DB 1; Length 868;
 Best Local Similarity 100.0%; Prod No. 0.75;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIVELLGRR 9
 |||||
 DB 792 RIVELLGRR 800

RESULT 13

P70175
 ID P70175 standard; Protein: 865 AA.
 AC P70175;
 DT 03-JUN-1991 (first entry)
 DE Sequence encoded by env gene of HTLV-III isolate HAT-3.
 KW Diagnosis; vaccine; HIV; LAV; gp 120; gp 41.
 OS HTLV-III.
 PN EP-227169-A.
 PD 01-JUL-1987.
 PE 08-DEC-1986; 202213.
 PR 17-DEC-1985; US-809872.
 PR 08-MAY-1986; US-861800.
 PR 18-JUL-1986; US-887294.
 PR 14-NOV-1986; US-927577.
 PA (ALKU) AKZO NV.
 PI Berman ML, Crush SA, Wong-Staal F, Gallo RC;
 DR WPI: 87-179470/26.

DR N-PSDB: N70259.
 PT Immunochemical reagent for HTLV-III - comprising 2 or more synthetic
 PI polypeptide sequences mimicking at least one antigenic determinant
 PT of gag antigen GP120 and GP41 of HTLV-III.
 PS Disclosure: Fig 3: 33pp; English.
 CC A novel immunochemical reagent which comprises a combination of 2
 CC or more synthetic polypeptide sequences selected from: the gag-
 CC antigen, glycoprotein gp 120 and glycoprotein gp 41 of HTLV-III is
 CC claimed. Also claimed are synthetic polypeptides comprising the
 CC carboxy terminal 38 AAs or acids 30-161 of the gp 120 protein, or
 CC AAs 53-122, 123-223 or 136-217 of the gp 41 protein. Also claimed
 CC are recombinant DNA and plasmids coding for the polypeptides.
 SQ Sequence 865 AA;

Query Match 100.0% Score 42. DB 1. Length 865;
 Best Local Similarity 100.0%; Pred. No. 0.75;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIVELLGRR 9
 |||||||
 DB 789 RIVELLGRR 797

RESULT 14
 ID P70665 standard; Protein: 901 AA.
 AC P70665:
 DT 23-APR-1991 (first entry)
 DE Sequence encoded by LAV/HTLV III envelope gene (env)
 KW Glycoprotein gp 120; gp 41. AIDS vaccine; diagnosis.
 OS LAV/HTLV III.
 PN W08702038-A.
 PD 09-APR-1987.
 PE 24-SEP-1985; 022987.
 PR 25-SEP-1985; US-779909.
 PR 27-MAR-1985; US-842984.
 PR 04-SEP-1986; US-905217.
 PA (ONCO-) ONCOGEN.
 PA (HUSC/) HU S.L.
 PI Hu Str. Purchio AF, Madisen L;
 DR WPI: 87-108683/15.
 DR N-PSDB: N71016.
 PI New recombinant viruses for directing expression of peptide(s)
 PI etc. - useful in vaccines for protecting humans against AIDS
 PI caused by LAV/HTLV III
 PS Claim 33: Fig 2: 165pp; English.
 CC Recombinant AC-NPV carrying the chimeric LAV/HTLV III env gene was
 CC used to infect Sf9 cells in tissue culture. The proteins produced on
 CC cultivation were immunoreactive with AIDS patient serum as well as
 CC with monoclonal antibodies which define LAV/HTLV III envelope
 CC glycoproteins gp. 110 and gp. 41. A recombinant DNA vector
 CC comprising ps-env 1,2,5 or 7 pv-gag1, pac-gag1 or pac-env 5, is
 CC claimed.
 SQ Sequence 901 AA;

Query Match 100.0% Score 42; DB 1; Length 901;
 Best Local Similarity 100.0%; Pred. No. 0.78;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIVELLGRR 9
 |||||||
 DB 825 RIVELLGRR 833

RESULT 15
 ID R12261 standard; Protein: 857 AA.
 AC R12261:
 DT 20-AUG-1991 (first entry)
 DE HIV-1 strain OY1 ENV Protein.
 KW HIV-1; AIDS; retroviruses.

OS Homo sapiens.
 PN US5019510-A.
 PD 28-MAY-1991.
 PE 28-OCT-1987; 113655.
 PR 28-OCT-1987; US-113655.
 PR (INSP) INST PASTEUR.
 PI Main-Hobson S, Huet T, Delaporte E, Brun-Vezinet F;
 DR WPI: 91-177518/24.
 PT Purified human retrovirus - is mutant of HIV-1 having
 PT characteristics of HIV-1 OY1, used in diagnosis of HIV infection
 PS Disclosure, Fig 7, 23pp; English.
 CC This sequence constitutes the ENV protein constituent of a new
 CC strain of HIV-1 retrovirus, OY1. This mutant retroviral strain is
 CC useful in an assay for diagnosing HIV infection. See also 011943
 CC (OY1 nucleotide sequence), P12275-60 and P12272 (other HIV OY1
 CC constituent proteins).
 SQ Sequence 857 AA;

Query Match 100.0% Score 42. DB 1. Length 857;
 Best Local Similarity 100.0%; Pred. No. 0.74;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIVELLGRR 9
 |||||||
 DB 781 RIVELLGRR 789

Search completed: December 9, 1999, 20:57:29
 Job time: 2165 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 1999, 22:13:49 : Search time 140.58 Seconds

(without alignments)
3.020 Million cell updates/sec

Title: US-09-017-524-33

Perfect score: 42

Sequence: 1 RIVELLGR 9

Scoring table: BLOSUM62

Searched: 142080 seqs; 4716939 residues

Database: PIR-52:
1: pir1:
2: pir2:
3: pir3:
4: pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB	ID	Description
1	42	100.0	856	1	VCLJH3	env polyprotein pr
2	42	100.0	856	1	VCLJVL	env polyprotein pr
3	42	100.0	851	1	VCLJTV	env polyprotein pr
4	42	100.0	843	1	H44001	env polyprotein pr
5	42	100.0	852	1	VCLJBR	env polyprotein pr
6	42	100.0	856	1	VCLJRW	env polyprotein pr
7	42	100.0	868	1	VCLJHA	env polyprotein pr
8	42	100.0	846	1	VCLJNB	env polyprotein pr
9	42	100.0	855	1	VCLJZR	env polyprotein pr
10	42	100.0	859	1	VCLJMN	env polyprotein pr
11	42	100.0	861	1	VCLJSC	env polyprotein pr
12	42	100.0	859	2	T31572	env polyprotein pr
13	42	100.0	357	2	S23005	env polyprotein pr
14	42	100.0	357	2	S21994	env polyprotein pr
15	42	100.0	358	2	S21998	env polyprotein pr
16	42	100.0	357	2	S21996	env polyprotein pr
17	42	100.0	357	2	S21996	env polyprotein pr
18	42	100.0	357	2	S21996	env polyprotein pr
19	42	100.0	357	2	S21996	env polyprotein pr
20	42	100.0	357	2	S21996	env polyprotein pr
21	42	100.0	357	2	S21996	env polyprotein pr
22	42	100.0	357	2	S21996	env polyprotein pr
23	42	100.0	357	2	S21996	env polyprotein pr
24	42	100.0	357	2	S21996	env polyprotein pr
25	42	100.0	357	2	S21996	env polyprotein pr
26	42	100.0	357	2	S21996	env polyprotein pr
27	42	100.0	357	2	S21996	env polyprotein pr
28	42	100.0	357	2	S21996	env polyprotein pr
29	42	100.0	357	2	S21996	env polyprotein pr
30	42	100.0	357	2	S21996	env polyprotein pr
31	42	100.0	357	2	S21996	env polyprotein pr
32	42	100.0	357	2	S21996	env polyprotein pr
33	42	100.0	357	2	S21996	env polyprotein pr
34	42	100.0	357	2	S21996	env polyprotein pr
35	42	100.0	357	2	S21996	env polyprotein pr
36	42	100.0	357	2	S21996	env polyprotein pr
37	42	100.0	357	2	S21996	env polyprotein pr
38	42	100.0	357	2	S21996	env polyprotein pr
39	42	100.0	357	2	S21996	env polyprotein pr

40 35 83.3 821.1 S76330 probable envelope
41 35 83.3 839.1 S71553 endopeptidase CIP
42 35 83.3 821.1 S73553 endopeptidase CIP
43 35 83.3 727.2 G02093 glycerol-3-phosphatase
44 34 81.0 861.1 VCLJBR env polyprotein pr
45 34 81.0 297.2 D44206 ribose-phosphate P

ALIGNMENTS

RESULT 1

VCLJH3

env polyprotein precursor : human immunodeficiency virus type 1 (isolate HTLV-III, BH

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 17-May-1985 #sequence_revision 17-May-1985 #rev_change 16-Jul-1999

C:Accession: A03973

P:Patel, L., Haseltine, W., Falarca, P., Livak, F.J., Staudh, E., Telford, S.F., D

berger, J.A., Papas, T.S., Graybe, J., Chang, N.T., Gallo, R.C., Wong-Staal, F.

Nature 313, 277-284, 1985

A:Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.

A:Reference number: A93553; M01D:8511123

A:Accession: A03973

A:Molecule type: DNA

A:Residues: 1-856 <RAT>

A:Cross-references: GB:M1554; GB:K00001; GB:K00002; GB:K00003; GB:K00004; GB:K00005; GB:K00006; GB:K00007; GB:K00008; GB:K00009; GB:K00010; GB:K00011; GB:K00012; GB:K00013; GB:K00014; GB:K00015; GB:K00016; GB:K00017; GB:K00018; GB:K00019; GB:K00020; GB:K00021; GB:K00022; GB:K00023; GB:K00024; GB:K00025; GB:K00026; GB:K00027; GB:K00028; GB:K00029; GB:K00030; GB:K00031; GB:K00032; GB:K00033; GB:K00034; GB:K00035; GB:K00036; GB:K00037; GB:K00038; GB:K00039; GB:K00040; GB:K00041; GB:K00042; GB:K00043; GB:K00044; GB:K00045; GB:K00046; GB:K00047; GB:K00048; GB:K00049; GB:K00050; GB:K00051; GB:K00052; GB:K00053; GB:K00054; GB:K00055; GB:K00056; GB:K00057; GB:K00058; GB:K00059; GB:K00060; GB:K00061; GB:K00062; GB:K00063; GB:K00064; GB:K00065; GB:K00066; GB:K00067; GB:K00068; GB:K00069; GB:K00070; GB:K00071; GB:K00072; GB:K00073; GB:K00074; GB:K00075; GB:K00076; GB:K00077; GB:K00078; GB:K00079; GB:K00080; GB:K00081; GB:K00082; GB:K00083; GB:K00084; GB:K00085; GB:K00086; GB:K00087; GB:K00088; GB:K00089; GB:K00090; GB:K00091; GB:K00092; GB:K00093; GB:K00094; GB:K00095; GB:K00096; GB:K00097; GB:K00098; GB:K00099; GB:K00100; GB:K00101; GB:K00102; GB:K00103; GB:K00104; GB:K00105; GB:K00106; GB:K00107; GB:K00108; GB:K00109; GB:K00110; GB:K00111; GB:K00112; GB:K00113; GB:K00114; GB:K00115; GB:K00116; GB:K00117; GB:K00118; GB:K00119; GB:K00120; GB:K00121; GB:K00122; GB:K00123; GB:K00124; GB:K00125; GB:K00126; GB:K00127; GB:K00128; GB:K00129; GB:K00130; GB:K00131; GB:K00132; GB:K00133; GB:K00134; GB:K00135; GB:K00136; GB:K00137; GB:K00138; GB:K00139; GB:K00140; GB:K00141; GB:K00142; GB:K00143; GB:K00144; GB:K00145; GB:K00146; GB:K00147; GB:K00148; GB:K00149; GB:K00150; GB:K00151; GB:K00152; GB:K00153; GB:K00154; GB:K00155; GB:K00156; GB:K00157; GB:K00158; GB:K00159; GB:K00160; GB:K00161; GB:K00162; GB:K00163; GB:K00164; GB:K00165; GB:K00166; GB:K00167; GB:K00168; GB:K00169; GB:K00170; GB:K00171; GB:K00172; GB:K00173; GB:K00174; GB:K00175; GB:K00176; GB:K00177; GB:K00178; GB:K00179; GB:K00180; GB:K00181; GB:K00182; GB:K00183; GB:K00184; GB:K00185; GB:K00186; GB:K00187; GB:K00188; GB:K00189; GB:K00190; GB:K00191; GB:K00192; GB:K00193; GB:K00194; GB:K00195; GB:K00196; GB:K00197; GB:K00198; GB:K00199; GB:K00200; GB:K00201; GB:K00202; GB:K00203; GB:K00204; GB:K00205; GB:K00206; GB:K00207; GB:K00208; GB:K00209; GB:K00210; GB:K00211; GB:K00212; GB:K00213; GB:K00214; GB:K00215; GB:K00216; GB:K00217; GB:K00218; GB:K00219; GB:K00220; GB:K00221; GB:K00222; GB:K00223; GB:K00224; GB:K00225; GB:K00226; GB:K00227; GB:K00228; GB:K00229; GB:K00230; GB:K00231; GB:K00232; GB:K00233; GB:K00234; GB:K00235; GB:K00236; GB:K00237; GB:K00238; GB:K00239; GB:K00240; GB:K00241; GB:K00242; GB:K00243; GB:K00244; GB:K00245; GB:K00246; GB:K00247; GB:K00248; GB:K00249; GB:K00250; GB:K00251; GB:K00252; GB:K00253; GB:K00254; GB:K00255; GB:K00256; GB:K00257; GB:K00258; GB:K00259; GB:K00260; GB:K00261; GB:K00262; GB:K00263; GB:K00264; GB:K00265; GB:K00266; GB:K00267; GB:K00268; GB:K00269; GB:K00270; GB:K00271; GB:K00272; GB:K00273; GB:K00274; GB:K00275; GB:K00276; GB:K00277; GB:K00278; GB:K00279; GB:K00280; GB:K00281; GB:K00282; GB:K00283; GB:K00284; GB:K00285; GB:K00286; GB:K00287; GB:K00288; GB:K00289; GB:K00290; GB:K00291; GB:K00292; GB:K00293; GB:K00294; GB:K00295; GB:K00296; GB:K00297; GB:K00298; GB:K00299; GB:K00300; GB:K00301; GB:K00302; GB:K00303; GB:K00304; GB:K00305; GB:K00306; GB:K00307; GB:K00308; GB:K00309; GB:K00310; GB:K00311; GB:K00312; GB:K00313; GB:K00314; GB:K00315; GB:K00316; GB:K00317; GB:K00318; GB:K00319; GB:K00320; GB:K00321; GB:K00322; GB:K00323; GB:K00324; GB:K00325; GB:K00326; GB:K00327; GB:K00328; GB:K00329; GB:K00330; GB:K00331; GB:K00332; GB:K00333; GB:K00334; GB:K00335; GB:K00336; GB:K00337; GB:K00338; GB:K00339; GB:K00340; GB:K00341; GB:K00342; GB:K00343; GB:K00344; GB:K00345; GB:K00346; GB:K00347; GB:K00348; GB:K00349; GB:K00350; GB:K00351; GB:K00352; GB:K00353; GB:K00354; GB:K00355; GB:K00356; GB:K00357; GB:K00358; GB:K00359; GB:K00360; GB:K00361; GB:K00362; GB:K00363; GB:K00364; GB:K00365; GB:K00366; GB:K00367; GB:K00368; GB:K00369; GB:K00370; GB:K00371; GB:K00372; GB:K00373; GB:K00374; GB:K00375; GB:K00376; GB:K00377; GB:K00378; GB:K00379; GB:K00380; GB:K00381; GB:K00382; GB:K00383; GB:K00384; GB:K00385; GB:K00386; GB:K00387; GB:K00388; GB:K00389; GB:K00390; GB:K00391; GB:K00392; GB:K00393; GB:K00394; GB:K00395; GB:K00396; GB:K00397; GB:K00398; GB:K00399; GB:K00400; GB:K00401; GB:K00402; GB:K00403; GB:K00404; GB:K00405; GB:K00406; GB:K00407; GB:K00408; GB:K00409; GB:K00410; GB:K00411; GB:K00412; GB:K00413; GB:K00414; GB:K00415; GB:K00416; GB:K00417; GB:K00418; GB:K00419; GB:K00420; GB:K00421; GB:K00422; GB:K00423; GB:K00424; GB:K00425; GB:K00426; GB:K00427; GB:K00428; GB:K00429; GB:K00430; GB:K00431; GB:K00432; GB:K00433; GB:K00434; GB:K00435; GB:K00436; GB:K00437; GB:K00438; GB:K00439; GB:K00440; GB:K00441; GB:K00442; GB:K00443; GB:K00444; GB:K00445; GB:K00446; GB:K00447; GB:K00448; GB:K00449; GB:K00450; GB:K00451; GB:K00452; GB:K00453; GB:K00454; GB:K00455; GB:K00456; GB:K00457; GB:K00458; GB:K00459; GB:K00460; GB:K00461; GB:K00462; GB:K00463; GB:K00464; GB:K00465; GB:K00466; GB:K00467; GB:K00468; GB:K00469; GB:K00470; GB:K00471; GB:K00472; GB:K00473; GB:K00474; GB:K00475; GB:K00476; GB:K00477; GB:K00478; GB:K00479; GB:K00480; GB:K00481; GB:K00482; GB:K00483; GB:K00484; GB:K00485; GB:K00486; GB:K00487; GB:K00488; GB:K00489; GB:K00490; GB:K00491; GB:K00492; GB:K00493; GB:K00494; GB:K00495; GB:K00496; GB:K00497; GB:K00498; GB:K00499; GB:K00500; GB:K00501; GB:K00502; GB:K00503; GB:K00504; GB:K00505; GB:K00506; GB:K00507; GB:K00508; GB:K00509; GB:K00510; GB:K00511; GB:K00512; GB:K00513; GB:K00514; GB:K00515; GB:K00516; GB:K00517; GB:K00518; GB:K00519; GB:K00520; GB:K00521; GB:K00522; GB:K00523; GB:K00524; GB:K00525; GB:K00526; GB:K00527; GB:K00528; GB:K00529; GB:K00530; GB:K00531; GB:K00532; GB:K00533; GB:K00534; GB:K00535; GB:K00536; GB:K00537; GB:K00538; GB:K00539; GB:K00540; GB:K00541; GB:K00542; GB:K00543; GB:K00544; GB:K00545; GB:K00546; GB:K00547; GB:K00548; GB:K00549; GB:K00550; GB:K00551; GB:K00552; GB:K00553; GB:K00554; GB:K00555; GB:K00556; GB:K00557; GB:K00558; GB:K00559; GB:K00560; GB:K00561; GB:K00562; GB:K00563; GB:K00564; GB:K00565; GB:K00566; GB:K00567; GB:K00568; GB:K00569; GB:K00570; GB:K00571; GB:K00572; GB:K00573; GB:K00574; GB:K00575; GB:K00576; GB:K00577; GB:K00578; GB:K00579; GB:K00580; GB:K00581; GB:K00582; GB:K00583; GB:K00584; GB:K00585; GB:K00586; GB:K00587; GB:K00588; GB:K00589; GB:K00590; GB:K00591; GB:K00592; GB:K00593; GB:K00594; GB:K00595; GB:K00596; GB:K00597; GB:K00598; GB:K00599; GB:K00600; GB:K00601; GB:K00602; GB:K00603; GB:K00604; GB:K00605; GB:K00606; GB:K00607; GB:K00608; GB:K00609; GB:K00610; GB:K00611; GB:K00612; GB:K00613; GB:K00614; GB:K00615; GB:K00616; GB:K00617; GB:K00618; GB:K00619; GB:K00620; GB:K00621; GB:K00622; GB:K00623; GB:K00624; GB:K00625; GB:K00626; GB:K00627; GB:K00628; GB:K00629; GB:K00630; GB:K00631; GB:K00632; GB:K00633; GB:K00634; GB:K00635; GB:K00636; GB:K00637; GB:K00638; GB:K00639; GB:K00640; GB:K00641; GB:K00642; GB:K00643; GB:K00644; GB:K00645; GB:K00646; GB:K00647; GB:K00648; GB:K00649; GB:K00650; GB:K00651; GB:K00652; GB:K00653; GB:K00654; GB:K00655; GB:K00656; GB:K00657; GB:K00658; GB:K00659; GB:K00660; GB:K00661; GB:K00662; GB:K00663; GB:K00664; GB:K00665; GB:K00666; GB:K00667; GB:K00668; GB:K00669; GB:K00670; GB:K00671; GB:K00672; GB:K00673; GB:K00674; GB:K00675; GB:K00676; GB:K00677; GB:K00678; GB:K00679; GB:K00680; GB:K00681; GB:K00682; GB:K00683; GB:K00684; GB:K00685; GB:K00686; GB:K00687; GB:K00688; GB:K00689; GB:K00690; GB:K00691; GB:K00692; GB:K00693; GB:K00694; GB:K00695; GB:K00696; GB:K00697; GB:K00698; GB:K00699; GB:K00700; GB:K00701; GB:K00702; GB:K00703; GB:K00704; GB:K00705; GB:K00706; GB:K00707; GB:K00708; GB:K00709; GB:K00710; GB:K00711; GB:K00712; GB:K00713; GB:K00714; GB:K00715; GB:K00716; GB:K00717; GB:K00718; GB:K00719; GB:K00720; GB:K00721; GB:K00722; GB:K00723; GB:K00724; GB:K00725; GB:K00726; GB:K00727; GB:K00728; GB:K00729; GB:K00730; GB:K00731; GB:K00732; GB:K00733; GB:K00734; GB:K00735; GB:K00736; GB:K00737; GB:K00738; GB:K00739; GB:K00740; GB:K00741; GB:K00742; GB:K00743; GB:K00744; GB:K00745; GB:K00746; GB:K00747; GB:K00748; GB:K00749; GB:K00750; GB:K00751; GB:K00752; GB:K00753; GB:K00754; GB:K00755; GB:K00756; GB:K00757; GB:K00758; GB:K00759; GB:K00760; GB:K00761; GB:K00762; GB:K00763; GB:K00764; GB:K00765; GB:K00766; GB:K00767; GB:K00768; GB:K00769; GB:K00770; GB:K00771; GB:K00772; GB:K00773; GB:K00774; GB:K00775; GB:K00776; GB:K00777; GB:K00778; GB:K00779; GB:K00780; GB:K00781; GB:K00782; GB:K00783; GB:K00784; GB:K00785; GB:K00786; GB:K00787; GB:K00788; GB:K00789; GB:K00790; GB:K00791; GB:K00792; GB:K00793; GB:K00794; GB:K00795; GB:K00796; GB:K00797; GB:K00798; GB:K00799; GB:K00800; GB:K00801; GB:K00802; GB:K00803; GB:K00804; GB:K00805; GB:K00806; GB:K00807; GB:K00808; GB:K00809; GB:K00810; GB:K00811; GB:K00812; GB:K00813; GB:K00814; GB:K00815; GB:K00816; GB:K00817; GB:K00818; GB:K00819; GB:K00820; GB:K00821; GB:K00822; GB:K00823; GB:K00824; GB:K00825; GB:K00826; GB:K00827; GB:K00828; GB:K00829; GB:K00830; GB:K00831; GB:K00832; GB:K00833; GB:K00834; GB:K00835; GB:K00836; GB:K00837; GB:K00838; GB:K00839; GB:K00840; GB:K00841; GB:K00842; GB:K00843; GB:K00844; GB:K00845; GB:K00846; GB:K00847; GB:K00848; GB:K00849; GB:K00850; GB:K00851; GB:K00852; GB:K00853; GB:K00854; GB:K00855; GB:K00856; GB:K00857; GB:K00858; GB:K00859; GB:K00860; GB:K00861; GB:K00862; GB:K00863; GB:K00864; GB:K00865; GB:K00866; GB:K00867; GB:K00868; GB:K00869; GB:K00870; GB:K00871; GB:K00872; GB:K00873; GB:K00874; GB:K00875; GB:K00876; GB:K00877; GB:K00878; GB:K00879; GB:K00880; GB:K00881; GB:K00882; GB:K00883; GB:K00884; GB:K00885; GB:K00886; GB:K00887; GB:K0

Query Match: 100.0%; Score 42; DB 1; Length 856;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 1 RIVELLGRR 9
 |||||
 DB 780 RIVELLGRR 788

RESULT 3
 VCLJLV
 env polyprotein precursor - human immunodeficiency virus type 1 (isolate LAV-1a)

N:Alternate names: coat polyprotein
 C:Species: human immunodeficiency virus type 1, HIV-1
 A:Note: host Homo sapiens (man)
 C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
 C:Accession: A04975
 R:Wain-Hobson, S.; Sonigo, P.; Barron, O.; Cole, S.; Allizon, M.
 Cell 40, 9-17, 1985
 A:Title: Nucleotide sequence of the AIDS virus, LAV.
 A:Reference number: A90866; MUID:85099333
 A:Accession: A04975
 A:Molecule type: DNA
 A:Residues: 1-861 <MAL>
 A:Cross-references: GR K02013; NID:9326417; PIDN AAB59751.1; PID:9326424
 C:Genetics:
 A:Gene: env
 C:Superfamily: type E retrovirus env polyprotein
 C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot
 F:1-30/Domain: signal sequence #status predicted <SIG>
 F:31-516/Product: exterior membrane glycoprotein #status predicted <EXT>
 F:517-861/Product: transmembrane glycoprotein #status predicted <TM>
 F:862-156,241,145,155,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411
 F:616,621,629,642,679,685,721/HindIII site: carbohydrate (Asn) (covalent) #status predic

Query Match: 100.0%; Score 42; DB 1; Length 861;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 1 RIVELLGRR 9
 |||||
 DB 785 RIVELLGRR 794

RESULT 4

env polyprotein precursor - human immunodeficiency virus type 1 (strain YU-2)

N:Contains: coat protein gp120; coat protein gp41
 C:Species: human immunodeficiency virus type 1, HIV-1
 A:Note: host Homo sapiens (man)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Sep-1994
 C:Accession: H44001
 R:Li, Y.; Hui, H.; Burgess, C. L.; Price, R. W.; Sharp, P. M.; Hahn, B. H.; Shaw, G. M.
 J Virol 66, 6567-6600, 1992
 A:Title: Complete nucleotide sequence, genome organization, and biological properties of
 A:Reference number: A44001; MUID:93021387
 A:Accession: H44001
 A:Molecule type: DNA
 A:Residues: 1-864 <LIV>
 A:Cross-references: GR M94258
 C:Genetics:
 A:Gene: env
 C:Superfamily: type E retrovirus env polyprotein
 C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:30-45/Region: hydrophobic
 F:46-489/Product: coat protein gp120 #status predicted <GP1>
 F:490-843/Product: coat protein gp41 #status predicted <GP2>
 F:849-855/Region: hydrophobic

F:673-689/Region: hydrophobic
 F:738-755/Domain: transmembrane #status predicted <TM>
 F:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,361,384,395,400,

Query Match: 100.0%; Score 42; DB 1; Length 843;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 1 RIVELLGRR 9
 |||||
 DB 767 RIVELLGRR 775

RESULT 5
 VCLJ3W
 env polyprotein - human immunodeficiency virus type 1 (isolate BR)

N:Alternate names: coat polyprotein
 N:Contains: coat protein gp120; coat protein gp41
 C:Species: human immunodeficiency virus type 1, HIV-1
 A:Note: host Homo sapiens (man)
 C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Feb-1997
 C:Accession: A31667
 R:Rand, R.; Thayer, P.; Srinivasan, A.; Nayyar, S.; Garbar, M.; Luciw, P.; Garbar
 Virology 168, 79-89, 1989
 A:Title: Biological and molecular characterization of human immunodeficiency virus (H
 A:Reference number: A94389; MUID:89085613
 A:Accession: A31667
 A:Molecule type: DNA
 A:Residues: 1-852 <ANA>
 C:Superfamily: type E retrovirus env polyprotein
 C:Keywords: capsid protein; coat protein; polyprotein; transmembrane protein
 F:1-516/Product: coat protein gp120 #status predicted <GP1>
 F:517-852/Product: coat protein gp41 #status predicted <GP2>

Query Match: 100.0%; Score 42; DB 1; Length 852;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 1 RIVELLGRR 9
 |||||
 DB 776 RIVELLGRR 784

RESULT 6

VCLJ3W
 env polyprotein precursor - human immunodeficiency virus type 1 (isolate WMJ1)

N:Contains: coat protein gp120; coat protein gp41
 C:Species: human immunodeficiency virus type 1, HIV-1
 C:Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 07-Nov-1997
 C:Accession: A24774
 R:Starick, B. R.; Hahn, B. H.; Shaw, G. M.; McNelly, P. D.; McDrew, S.; Wolf, H.; Parks,
 Cell 45, 637-648, 1986
 A:Title: Identification and characterization of conserved and variable regions in the
 A:Reference number: A24774; MUID:86218077
 A:Accession: A24774
 A:Molecule type: DNA
 A:Residues: 1-856 <STA>
 A:Cross-references: GR K03455; GR M8432; NID:9305382
 C:Genetics:
 A:Gene: env
 C:Superfamily: type E retrovirus env polyprotein
 C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:30-501/Product: coat protein gp120 #status predicted <GP1>
 F:502-847/Product: coat protein gp41 #status predicted <GP2>
 F:87,134,140,151,155,183,197,234,241,262,276,289,295,311,338,364,370,390,394,401,447,

Query Match: 100.0%; Score 42; DB 1; Length 856;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIVELLGRR 9
 |||||||
 Db 780 RIVELLGRR 788

RESULT 7

VCLJH4
 env polyprotein precursor - human immunodeficiency virus type 1 (isolate CDC-451)

N:Alternate names: coat polyprotein

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999

C:Accession: D25523

P:Desai S M.; Kalyanaraman V S.; Casey J M.; Srinivasan A.; Andersen P.R.; Devare, Proc. Natl. Acad. Sci. U.S.A. 83, 8380-8384, 1986

A:Title: Molecular cloning and primary nucleotide sequence analysis of a distinct human A:Reference number: A94136, MIMD:87041461

A:Accession: D25523

A:Molecule type: DNA

A:Residues: 1-866 <SR>

A:Cross-references: GR:M11137, NID:9326460, PIDN:AAA4411.1, PID:9326467

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane prote

F:1-521/Product: coat protein gp120 #status predicted <CP1>

F:522-866/Product: coat protein gp41 #status predicted <CP2>

F:867-1441/Product: coat protein gp120 #status predicted <CP3>

F:1442-1447/Product: coat protein gp120 #status predicted <CP4>

F:1448-1453/Product: coat protein gp120 #status predicted <CP5>

F:1454-1459/Product: coat protein gp120 #status predicted <CP6>

F:1460-1465/Product: coat protein gp120 #status predicted <CP7>

F:1466-1471/Product: coat protein gp120 #status predicted <CP8>

F:1472-1477/Product: coat protein gp120 #status predicted <CP9>

F:1478-1483/Product: coat protein gp120 #status predicted <CP10>

F:1484-1489/Product: coat protein gp120 #status predicted <CP11>

F:1490-1495/Product: coat protein gp120 #status predicted <CP12>

F:1496-1501/Product: coat protein gp120 #status predicted <CP13>

F:1502-1507/Product: coat protein gp120 #status predicted <CP14>

F:1508-1513/Product: coat protein gp120 #status predicted <CP15>

F:1514-1519/Product: coat protein gp120 #status predicted <CP16>

F:1520-1525/Product: coat protein gp120 #status predicted <CP17>

F:1526-1531/Product: coat protein gp120 #status predicted <CP18>

F:1532-1537/Product: coat protein gp120 #status predicted <CP19>

F:1538-1543/Product: coat protein gp120 #status predicted <CP20>

F:1544-1549/Product: coat protein gp120 #status predicted <CP21>

F:1550-1555/Product: coat protein gp120 #status predicted <CP22>

F:1556-1561/Product: coat protein gp120 #status predicted <CP23>

F:1562-1567/Product: coat protein gp120 #status predicted <CP24>

F:1568-1573/Product: coat protein gp120 #status predicted <CP25>

F:1574-1579/Product: coat protein gp120 #status predicted <CP26>

F:1580-1585/Product: coat protein gp120 #status predicted <CP27>

F:1586-1591/Product: coat protein gp120 #status predicted <CP28>

F:1592-1597/Product: coat protein gp120 #status predicted <CP29>

F:1598-1603/Product: coat protein gp120 #status predicted <CP30>

F:1604-1609/Product: coat protein gp120 #status predicted <CP31>

F:1610-1615/Product: coat protein gp120 #status predicted <CP32>

F:1616-1621/Product: coat protein gp120 #status predicted <CP33>

F:1622-1627/Product: coat protein gp120 #status predicted <CP34>

F:1628-1633/Product: coat protein gp120 #status predicted <CP35>

F:1634-1639/Product: coat protein gp120 #status predicted <CP36>

F:1640-1645/Product: coat protein gp120 #status predicted <CP37>

F:1646-1651/Product: coat protein gp120 #status predicted <CP38>

QY 1 RIVELLGRR 9
 |||||||
 Db 770 RIVELLGRR 778

RESULT 9

VCLJH4
 env polyprotein precursor - human immunodeficiency virus Zr-6

N:Alternate names: coat polyprotein

C:Species: human immunodeficiency virus Zr-6

C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999

C:Accession: D26192

P:Srinivasan A.; Anand R.; York D.; Fungnathan, F.; Fungnathan, P.; Scholten, G.; Gene 52, 71-82, 1987

A:Title: Molecular characterization of human immunodeficiency virus from Zaire: nucle A:Reference number: A26192, MIMD:87248097

A:Accession: D26192

A:Molecule type: DNA

A:Residues: 1-855 <SR>

A:Cross-references: GR:M05458, GR:M11137, NID:9326460, PIDN:AAA4411.1, PID:9326467

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane prote

F:1-19/Product: signal sequence #status predicted <SIS>

F:20-85/Product: env polyprotein #status predicted <WAT>

F:86-500/Product: exterior membrane glycoprotein #status predicted <EXT>

F:501-855/Product: transmembrane glycoprotein #status predicted <TM2>

F:856-1400/Product: transmembrane glycoprotein #status predicted <TM3>

F:1401-1406/Product: transmembrane glycoprotein #status predicted <TM4>

F:1407-1412/Product: transmembrane glycoprotein #status predicted <TM5>

F:1413-1418/Product: transmembrane glycoprotein #status predicted <TM6>

F:1419-1424/Product: transmembrane glycoprotein #status predicted <TM7>

F:1425-1430/Product: transmembrane glycoprotein #status predicted <TM8>

F:1431-1436/Product: transmembrane glycoprotein #status predicted <TM9>

F:1437-1442/Product: transmembrane glycoprotein #status predicted <TM10>

F:1443-1448/Product: transmembrane glycoprotein #status predicted <TM11>

F:1449-1454/Product: transmembrane glycoprotein #status predicted <TM12>

F:1455-1460/Product: transmembrane glycoprotein #status predicted <TM13>

F:1461-1466/Product: transmembrane glycoprotein #status predicted <TM14>

F:1467-1472/Product: transmembrane glycoprotein #status predicted <TM15>

F:1473-1478/Product: transmembrane glycoprotein #status predicted <TM16>

F:1479-1484/Product: transmembrane glycoprotein #status predicted <TM17>

F:1485-1490/Product: transmembrane glycoprotein #status predicted <TM18>

F:1491-1496/Product: transmembrane glycoprotein #status predicted <TM19>

F:1497-1502/Product: transmembrane glycoprotein #status predicted <TM20>

F:1503-1508/Product: transmembrane glycoprotein #status predicted <TM21>

F:1509-1514/Product: transmembrane glycoprotein #status predicted <TM22>

F:1515-1520/Product: transmembrane glycoprotein #status predicted <TM23>

F:1521-1526/Product: transmembrane glycoprotein #status predicted <TM24>

F:1527-1532/Product: transmembrane glycoprotein #status predicted <TM25>

F:1533-1538/Product: transmembrane glycoprotein #status predicted <TM26>

F:1539-1544/Product: transmembrane glycoprotein #status predicted <TM27>

F:1545-1550/Product: transmembrane glycoprotein #status predicted <TM28>

F:1551-1556/Product: transmembrane glycoprotein #status predicted <TM29>

F:1557-1562/Product: transmembrane glycoprotein #status predicted <TM30>

F:1563-1568/Product: transmembrane glycoprotein #status predicted <TM31>

F:1569-1574/Product: transmembrane glycoprotein #status predicted <TM32>

F:1575-1580/Product: transmembrane glycoprotein #status predicted <TM33>

F:1581-1586/Product: transmembrane glycoprotein #status predicted <TM34>

F:1587-1592/Product: transmembrane glycoprotein #status predicted <TM35>

F:1593-1598/Product: transmembrane glycoprotein #status predicted <TM36>

QY 1 RIVELLGRR 9
 |||||||
 Db 783 RIVELLGRR 791

RESULT 10

VCLJH4
 env polyprotein precursor - human immunodeficiency virus type 1 (isolate MN)

N:Alternate names: coat polyprotein

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Feb-1997

C:Accession: A28922

P:Gurgu, G.; Guo, H.C.; Franchini, G.; Albovicio, A.; Callaway, E.; Farrell, K.; Wong- R:Gurgu, G.; Guo, H.C.; Franchini, G.; Albovicio, A.; Callaway, E.; Farrell, K.; Wong- A:Title: Envelope sequences of two new United States HIV-1 isolates.

A:Reference number: A28922, MIMD:86219542

A:Accession: A28922

A:Molecule type: DNA

A:Residues: 1-859 <GUR>

A:Cross-references: GR:M05458, GR:M11137, NID:9326460, PIDN:AAA4411.1, PID:9326467

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane pr

F:1-19/Product: signal sequence #status predicted <SIS>

F:20-85/Product: env polyprotein #status predicted <WAT>

F:86-500/Product: exterior membrane glycoprotein #status predicted <EXT>

F:501-855/Product: transmembrane glycoprotein #status predicted <TM2>

F:856-1400/Product: transmembrane glycoprotein #status predicted <TM3>

F:1401-1406/Product: transmembrane glycoprotein #status predicted <TM4>

F:1407-1412/Product: transmembrane glycoprotein #status predicted <TM5>

F:1413-1418/Product: transmembrane glycoprotein #status predicted <TM6>

F:1419-1424/Product: transmembrane glycoprotein #status predicted <TM7>

F:1425-1430/Product: transmembrane glycoprotein #status predicted <TM8>

F:1431-1436/Product: transmembrane glycoprotein #status predicted <TM9>

F:1437-1442/Product: transmembrane glycoprotein #status predicted <TM10>

F:1443-1448/Product: transmembrane glycoprotein #status predicted <TM11>

F:1449-1454/Product: transmembrane glycoprotein #status predicted <TM12>

F:1455-1460/Product: transmembrane glycoprotein #status predicted <TM13>

F:1461-1466/Product: transmembrane glycoprotein #status predicted <TM14>

F:1467-1472/Product: transmembrane glycoprotein #status predicted <TM15>

F:1473-1478/Product: transmembrane glycoprotein #status predicted <TM16>

F:1479-1484/Product: transmembrane glycoprotein #status predicted <TM17>

F:1485-1490/Product: transmembrane glycoprotein #status predicted <TM18>

F:1491-1496/Product: transmembrane glycoprotein #status predicted <TM19>

F:1497-1502/Product: transmembrane glycoprotein #status predicted <TM20>

F:1503-1508/Product: transmembrane glycoprotein #status predicted <TM21>

F:1509-1514/Product: transmembrane glycoprotein #status predicted <TM22>

F:1515-1520/Product: transmembrane glycoprotein #status predicted <TM23>

F:1521-1526/Product: transmembrane glycoprotein #status predicted <TM24>

F:1527-1532/Product: transmembrane glycoprotein #status predicted <TM25>

F:1533-1538/Product: transmembrane glycoprotein #status predicted <TM26>

F:1539-1544/Product: transmembrane glycoprotein #status predicted <TM27>

F:1545-1550/Product: transmembrane glycoprotein #status predicted <TM28>

F:1551-1556/Product: transmembrane glycoprotein #status predicted <TM29>

F:1557-1562/Product: transmembrane glycoprotein #status predicted <TM30>

F:1563-1568/Product: transmembrane glycoprotein #status predicted <TM31>

F:1569-1574/Product: transmembrane glycoprotein #status predicted <TM32>

F:1575-1580/Product: transmembrane glycoprotein #status predicted <TM33>

F:1581-1586/Product: transmembrane glycoprotein #status predicted <TM34>

F:1587-1592/Product: transmembrane glycoprotein #status predicted <TM35>

F:1593-1598/Product: transmembrane glycoprotein #status predicted <TM36>

Wed Dec 15 08:56:50 1999

us-09-017-524-33.rpr

Page 5

Job time: 1510 sec

1

1

1